

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
8 March 2001 (08.03.2001)

PCT

(10) International Publication Number
WO 01/16340 A1

(51) International Patent Classification⁷: C12N 15/82,
A01H 5/10, 5/00, C12N 15/29

Cove, N.W., Calgary, Alberta T3A 5N5 (CA). SINGH,
Surinder [AU/AU]; 10 Lucas Place, Downer, ACT 2601
(AU).

(21) International Application Number: PCT/CA00/00988

(74) Agent: BERESKIN & PARR; 40 King Street West, 40th
Floor, Toronto, Ontario M5H 3Y2 (CA).

(22) International Filing Date: 25 August 2000 (25.08.2000)

(25) Filing Language: English

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(26) Publication Language: English

(30) Priority Data:
60/151,044 27 August 1999 (27.08.1999) US
60/161,722 27 October 1999 (27.10.1999) US
2,310,304 30 May 2000 (30.05.2000) CA

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicants (*for all designated States except US*): SEM-
BIOSYS GENETICS INC. [CA/CA]; 2500 University
Avenue Drive N.W., Room 460, Calgary, Alberta T2N 1N4
(CA). COMMONWEALTH SCIENTIFIC AND IN-
DUSTRIAL RESEARCH ORGANISATION [AU/AU];
CSIRO Plant Industry, GPO Box 1600, Canberra, ACT
2601 (AU).

Published:

- With international search report.
- Before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments.

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): CHAUDHARY,
Sarita [CA/CA]; 193 Hawkbury Close N.W., Calgary,
Alberta T3G 3C9 (CA). VAN ROOIJEN, Gijs [NL/CA];
3223 Bears paw Drive N.W., Calgary, Alberta T2L 1T1
(CA). MOLONEY, Maurice, M. [IE/CA]; 34 Edgebrook

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: FLAX SEED SPECIFIC PROMOTERS

(57) Abstract: Novel methods for the expression of non-native genes in flax seeds and the seeds of other plant species are provided. The methods involve the use of seed-specific promoters obtained from flax. Additionally provided are novel flax seed-specific promoters, chimeric nucleic acid constructs comprising novel flax seed-specific promoters, transgenic plant cells, transgenic plants and transgenic plant seeds containing novel flax seed-specific promoters. The promoters and methods are useful, for example, for altering the seed oil and protein composition in flax seed or other plant seeds.

WO 01/16340 A1

Title: Flax Seed Specific Promoters

FIELD OF THE INVENTION

The present invention relates to plant genetic engineering methods useful for the alteration of the constituents of plant seeds. More specifically, the invention relates to promoters that have been obtained from flax and are capable of directing expression of non-native genes in flax seeds as well as the seeds of other plants.

BACKGROUND OF THE INVENTION

Flax or linseed (*Linum usitatissimum*) is a commercially important oilseed crop. Flax oil and meal are valuable raw materials derived from flax seed. A further economically significant raw material, flax fiber, is obtainable from the stem of the plant. The flax oil fraction is used for non-edible purposes, for example in the manufacture of varnish and paint, and has more recently become suited for use in the manufacture of a range of edible products, such as margarines and salad oils and dressings, by virtue of newly bred so called Linola cultivars (Green (1986) Can. J. Plant Sci, 66: 499-503). Flax meal is used primarily as a constituent of ruminant feeds while flax fibers are used in the manufacture of linen fabrics. Given its economic importance as a source for raw materials, it is desirable to further improve and diversify the available flax cultivar portfolio both with respect to agronomic performance, for example seed yield, resistance to pathogens and low climatic temperatures, and with respect to yield and quality of the raw materials to suit downstream applications. Although it is possible to obtain improved flax cultivars through conventional plant breeding, as evidenced by the development of the Linola cultivars, developing an elite agronomic plant line requires large investments in plant breeding due to the long timelines involved. Plant genetic engineering technology allows the isolation of genes directly from unrelated species and the transfer of these genes into elite agronomic backgrounds, thereby significantly reducing the time required to develop new cultivars. In addition plant genetic engineering permits the manufacture of products not naturally obtainable from flax, for example therapeutic agents.

In order to develop novel flax cultivars through plant genetic engineering, control over the expression of the introduced foreign or non-native gene is of critical importance. The desired expression characteristics for the non-native gene, such as the level of expression of the non-native gene, the particular plant tissue or organ in which the non-native gene is expressed, and the particular time in the growth cycle of the plant at which the non-native gene is expressed, will vary depending on the application for which the plant line is developed. For example, the modification of the seed oil composition may require low levels of seed-specific expression of an enzyme involved in fatty acid metabolism at an early stage in seed development (see for example US Patent 5,420,034). On the other hand expression of a pharmaceutical protein could preferably require high

levels of leaf-specific expression upon harvesting of the plant leaves (see for example, US Patent 5,929,304).

In order to manipulate the expression characteristics of non-native genes numerous factors can be influenced. One factor is the choice of the transcriptional promoter used. A wide range of plant compatible promoters is currently available and some of the better documented promoters include constitutive promoters such as the 35-S CaMV promoter (Rothstein *et al.* (1987), Gene 53: 153-161) and the ubiquitin promoter (US Patent 5,614,399), tissue specific promoters such as seed-specific promoters, for example the phaseolin promoter (Sengupta-Gopalan *et al.*, (1985), PNAS USA 82: 3320-3324) and inducible promoters, such as those inducible by heat (Czarnencka *et al.*, (1989), Mol. Cell. Biol. 9 (8): 3457-3464), UV light, elicitors and wounding (Lois *et al.*, (1989) EMBO J. 8 (6): 1641-1648), or chemicals such as endogenous hormones (Skriver *et al.* (1991), Proc. Natl. Acad. Sci. USA 88(16): 7266-7270). Other factors that can be manipulated in order to control the expression characteristics of non-native gene in plants include transcriptional modification factors such as introns, polyadenylation sites and transcription termination sites. The expression characteristics of the non-native gene can further be manipulated by factors that affect translation, such as ribosomal binding sites and the codon bias that is exhibited by the host. Furthermore, the non-native gene itself may affect the viability of the transgenic plant, thus limiting particularly the levels of expression that can be attained. In some cases it may be possible to overcome this problem, by expressing the protein in a tissue specific manner, e.g. in the leaves or seed, or by restricting the accumulation of the protein in different subcellular compartments such as for example the cytoplasm, the endoplasmic reticulum or vacuoles, typically by the presence or the absence of specific targeting sequences capable of directing the protein to these compartments. Another factor that will affect the expression characteristics is the location in which the construct inserts itself into the host chromosome. This effect could provide an explanation as to why different plants, transformed with the same recombinant construct, can have fluctuating levels of recombinant protein expression.

To the best of the inventors' knowledge, expression of non-native genes in flax seeds is only documented in PCT Patent Application WO 98/18948. This application discloses two stearyl-acyl carrier protein desaturase (SAD) genes derived from flax. The associated SAD promoter sequences are useful for the modification of flax and other plants for the expression of endogenous or foreign genes. However the methods taught by WO 98/18948 are limited by the fact that the SAD promoters are not seed-specific in flax and confer expression to leaves, stems, flowers and seeds. Expression of non-native genes thus may result in undesirable side effects in non-seed tissues. In addition the use of the SAD promoters allows limited control over expression level and timing of expression.

There is a need in the art to further improve methods for the expression of non-native genes in flax seeds and other plant seeds.

SUMMARY OF THE INVENTION

The present invention relates to improved methods for the seed-specific
5 expression of non-native genes in plants. In particular, the invention relates to improved methods for the seed-specific expression of non-native genes in flax.

Accordingly, in one aspect, the invention provides a method for the expression of a nucleic acid sequence of interest in flax seeds comprising:

- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3'
10 direction of transcription as operably linked components
 - (1) a seed-specific promoter obtained from flax; and
 - (2) the nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said flax seed-specific promoter;
- (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
- 15 (c) growing said flax plant cell into a mature flax plant capable of setting seed, wherein said nucleic acid sequence of interest is expressed in the seed under the control of said flax seed-specific promoter.

In a preferred embodiment of the invention, at least one expression characteristic, e.g. timing of expression in the plant's life cycle, conferred by the promoter to the
20 non-native nucleic acid sequence is similar to that expression characteristic when conferred to a native nucleic acid sequence. In further preferred embodiments, the flax seed-specific promoter is an oleosin promoter, a 2S storage protein promoter or a legumin-like seed storage protein promoter.

In a further aspect, the present invention provides transgenic flax seeds prepared
25 according to a method comprising:

- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3'
direction of transcription as operably linked components:
 - (1) a seed-specific promoter obtained from flax; and
 - (2) a nucleic acid sequence of interest wherein said nucleic acid of interest is
30 non-native to said seed-specific promoter;
- (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
- (c) growing said flax plant cell into a mature flax plant capable of setting seed, wherein said nucleic acid sequence of interest is expressed in the seed under the control of said seed-specific promoter.

In a further aspect the present invention provides flax plants capable of setting
35 seed prepared by a method comprising:

- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3'
direction of transcription as operably linked components:

- (1) a seed-specific promoter obtained from flax; and
- (2) a nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said seed-specific promoter;
- (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
- 5 (c) growing said flax plant cell into a mature flax plant capable of setting seed, wherein said nucleic acid sequence of interest is expressed in the seed under the control of said seed-specific promoter.

In yet a further aspect, the present invention provides novel flax seed specific promoters useful for the expression of non-native genes in flax seeds and the seeds of other
 10 plant species, useful for example for modification of the protein or oil composition of the seed.

In a preferred embodiment, the seed specific promoter comprises:

- (a) a nucleic acid sequence as shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8) wherein T can also be
 15 U;
- (b) a nucleic acid sequence that is complimentary to a nucleic acid sequence of (a);
- (c) a nucleic acid sequence that has substantial sequence homology to a nucleic acid sequence of (a) or (b);
- 20 (d) a nucleic acid sequence that is an analog of a nucleic acid sequence of (a), (b) or (c); or
- (e) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (a), (b), (c) or (d) under stringent hybridization conditions.

In another aspect, the invention provides chimeric nucleic acid sequences
 25 comprising a first nucleic acid sequence obtained from flax operatively linked to a second nucleic acid sequence non-native to said first nucleic acid sequence wherein said first nucleic acid sequence comprises a novel flax seed-specific promoter.

Other features and advantages of the present invention will become readily apparent from the following detailed description. It should be understood, however, that
 30 the detailed description and the specific examples while indicating preferred embodiments of the invention are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art of this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

35 The invention will now be described in relation to the drawings in which:

Figure 1 shows the DNA sequence (SEQ.ID.NO.:1) of a flax genomic clone encoding a 16.0 kDa oleosin protein (SEQ.ID.NOS.:2 and 3).

Figure 2 shows the DNA sequence (SEQ.ID.NO.:4) of a flax genomic clone encoding a 18.6 kDa oleosin protein (SEQ.ID.NO.:5).

Figure 3 shows the DNA sequence (SEQ.ID.NO.:6) of a flax genomic clone encoding a 2S storage protein (SEQ.ID.NO.:7).

5 Figure 4 shows the DNA sequence (SEQ.ID.NO.:8) of a flax genomic clone encoding a 54.5 kDa legumin-like storage protein (SEQ.ID.NOS.:9-12).

Figure 5 shows Southern blot analysis of flax genomic DNA probed with flax oleosin DNA sequences.

10 Figure 6 shows a Northern blot analysis of the seed specific expression of flax oleosins.

Figure 7 shows a Northern blot analysis of the developmental expression of flax oleosins during seed development.

Figure 8 shows the GUS activity of flax embryos bombarded with flax oleosin promoter-GUS-flax terminator gene constructs.

15 Figure 9 shows GUS expression in developing flax embryos and Arabidopsis seeds of plants transformed with a 2S protein gene promoter GUS fusion.

Figure 10 shows the tissue-specific expression of GUS in transgenic flax plants transformed with a linin promoter-GUS-linin terminator gene construct.

20 Figure 11 shows the temporal expression of GUS in transgenic flax plants transformed a linin promoter-GUS-linin terminator gene construct.

Figure 12 shows the expression of GUS in transgenic *Brassica napus* plants (L1 to L9) transformed with a linin promoter-GUS-linin terminator gene construct.

25 Figure 13 shows the expression of GUS in transgenic Arabidopsis plants transformed with a linin promoter-GUS-linin terminator gene construct at different stages of seed development.

DETAILED DESCRIPTION OF THE INVENTION

As hereinbefore mentioned, the present invention relates to improved methods for the expression of non-native genes in plants, in particular flax. The invention provides methods allowing the seed-specific expression of non-native genes in flax. The methods of
30 the invention are advantageous in that improved control over the expression of non-native genes in flax seeds is obtained. Expression of the non-native gene is restricted to the seeds, thereby limiting potential undesirable effects resulting from the expression in other plant organs or tissues. In addition, the provided methodology allows improved control over expression characteristics, such as the expression level of the non-native gene and timing of
35 expression of the non-native gene in the developmental cycle of the plant. The methods of the present invention are particularly useful in that in accordance with the present invention the seed composition with respect to valuable raw materials, such as oil, protein and polysaccharides, may be altered both qualitatively and quantitatively.

Accordingly, in one aspect, the invention provides a method for the expression of a nucleic acid sequence of interest in flax seeds comprising:

- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3' direction of transcription as operably linked components;
 - 5 (1) a seed-specific promoter obtained from flax; and
 - (2) the nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said flax seed-specific promoter;
 - (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
 - (c) growing said flax plant cell into a mature plant capable of setting seed,
- 10 wherein said nucleic acid sequence of interest is expressed in the seed under the control of said flax seed-specific promoter.

As used herein, the term "non-native" refers to any nucleic acid sequence, including any RNA or DNA sequence, which is not normally associated with the seed-specific promoter. This includes heterologous nucleic acid sequences which are
15 obtained from a different plant species as the promoter as well as homologous nucleic acid sequences which are obtained from the same plant species as the promoter but are not associated with the promoter in the wild-type (non-transgenic) plant.

The non-native nucleic acid sequence when linked to a seed-specific promoter obtained from flax results in a chimeric construct. The chimeric construct is introduced into a
20 flax plant cell to create a transgenic flax plant cell which results in a detectably different phenotype of the flax plant cell or flax plant grown from it when compared with a non-transgenic flax plant cell or flax plant grown from it. A contiguous nucleic acid sequence identical to the nucleic acid sequence of the chimeric construct is not present in the non-transformed flax plant cell or flax plant grown from it. In this respect, chimeric nucleic
25 acid sequences include those sequences which contain a flax promoter linked to a nucleic acid sequence obtained from another plant species or a nucleic acid sequence from flax but normally not associated with that promoter. Chimeric nucleic acid sequences as used herein further include sequences comprising a flax promoter and a nucleic acid sequence that is normally linked to the promoter but additionally containing a non-native nucleic acid
30 sequence. For example, if the promoter is a flax seed-specific oleosin promoter, sequences "non-native" to the flax oleosin promoter also include a sequence comprising a fusion between the flax oleosin gene naturally associated with the oleosin promoter, and a coding sequence of interest that is not naturally associated with the promoter. The term non-native is also meant to include a fusion gene as hereinabove which additionally
35 includes a cleavage sequence separating the nucleic acid sequence that is normally linked to the promoter sequence and the gene encoding the protein of interest.

The term "nucleic acid sequence" refers to a sequence of nucleotide or nucleoside monomers consisting of naturally occurring bases, sugars and intersugar (backbone) linkages.

The term also includes modified or substituted sequences comprising non-naturally occurring monomers or portions thereof, which function similarly. The nucleic acid sequences of the present invention may be ribonucleic (RNA) or deoxyribonucleic acids (DNA) and may contain naturally occurring bases including adenine, guanine, cytosine, thymidine and uracil. The sequences may also contain modified bases such as xanthine, hypoxanthine, 2-aminoadenine, 6-methyl, 2-propyl, and other alkyl adenines, 5-halo uracil, 5-halo cytosine, 6-aza uracil, 6-aza cytosine and 6-aza thymine, pseudo uracil, 4-thiouracil, 8-halo adenine, 8-amino adenine, 8-thiol adenine, 8-thio-alkyl adenines, 8-hydroxyl adenine and other 8-substituted adenines, 8-halo guanines, 8-amino guanine, 8-thiol guanine, 8-thioalkyl guanines, 8-hydroxyl guanine and other 8-substituted guanines, other aza and deaza uracils, thymidines, cytosines, adenines, or guanines, 5-trifluoromethyl uracil and 5-trifluoro cytosine.

The term "seed-specific promoter", means that a gene expressed under the control of the promoter is predominantly expressed in plant seeds with no or no substantial expression, typically less than 5% of the overall expression level, in other plant tissues.

In a further aspect, the present invention provides novel flax seed specific promoters useful for the expression of non-native genes in flax seeds and the seeds of other plant species. The promoters may be used to modify for example the protein, oil or polysaccharide composition of the seeds. In a preferred embodiment, the seed specific promoter comprises:

- (a) a nucleic acid sequence as shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8) wherein T can also be U;
- (b) a nucleic acid sequence that is complimentary to a nucleic acid sequence of (a);
- (c) a nucleic acid sequence that has substantial sequence homology to a nucleic acid sequence of (a) or (b);
- (d) a nucleic acid sequence that is an analog of a nucleic acid sequence of (a), (b) or (c); or
- (e) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (a), (b), (c) or (d) under stringent hybridization conditions.

The term "sequence that has substantial sequence homology" means those nucleic acid sequences which have slight or inconsequential sequence variations from the sequences in (a) or (b), i.e., the sequences function in substantially the same manner and are capable of driving seed specific expression of non-native nucleic acid sequences. The variations may be attributable to local mutations or structural modifications. Nucleic acid sequences having substantial homology include nucleic acid sequences having at least 65%, more preferably at least 85%, and most preferably 90-95% identity with the nucleic acid sequences as shown in

Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8).

The term "sequence that hybridizes" means a nucleic acid sequence that can hybridize to a sequence of (a), (b), (c) or (d) under stringent hybridization conditions.

5 Appropriate "stringent hybridization conditions" which promote DNA hybridization are known to those skilled in the art, or may be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the following may be employed: 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C. The stringency may be selected based on the conditions used in the wash step. For
10 example, the salt concentration in the wash step can be selected from a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be at high stringency conditions, at about 65°C.

The term "a nucleic acid sequence which is an analog" means a nucleic acid sequence which has been modified as compared to the sequence of (a), (b) or (c) wherein the
15 modification does not alter the utility of the sequence (i.e. as a seed specific promoter) as described herein. The modified sequence or analog may have improved properties over the sequence shown in (a), (b) or (c). One example of a modification to prepare an analog is to replace one of the naturally occurring bases (i.e. adenine, guanine, cytosine or thymidine) of the sequence shown in Figure 1, Figure 2, Figure 3 or Figure 4 with a modified base such as
20 such as xanthine, hypoxanthine, 2-aminoadenine, 6-methyl, 2-propyl and other alkyl adenines, 5-halo uracil, 5-halo cytosine, 6-aza uracil, 6-aza cytosine and 6-aza thymine, pseudo uracil, 4-thiouracil, 8-halo adenine, 8-aminoadenine, 8-thiol adenine, 8-thiolalkyl adenines, 8-hydroxyl adenine and other 8-substituted adenines, 8-halo guanines, 8 amino guanine, 8-thiol guanine, 8-thiolalkyl guanines, 8-hydroxyl guanine and other
25 8-substituted guanines, other aza and deaza uracils, thymidines, cytosines, adenines, or guanines, 5-trifluoromethyl uracil and 5-trifluoro cytosine.

Another example of a modification is to include modified phosphorous or oxygen heteroatoms in the phosphate backbone, short chain alkyl or cycloalkyl intersugar linkages or short chain heteroatomic or heterocyclic intersugar linkages in the nucleic acid
30 molecule shown in Figure 1, Figure 2, Figure 3 or Figure 4. For example, the nucleic acid sequences may contain phosphorothioates, phosphotriesters, methyl phosphonates, and phosphorodithioates.

A further example of an analog of a nucleic acid molecule of the invention is a peptide nucleic acid (PNA) wherein the deoxyribose (or ribose) phosphate backbone in the
35 DNA (or RNA), is replaced with a polyamide backbone which is similar to that found in peptides (P.E. Nielsen, et al Science 1991, 254, 1497). PNA analogs have been shown to be resistant to degradation by enzymes and to have extended lives *in vivo* and *in vitro*. PNAs also bind stronger to a complimentary DNA sequence due to the lack of charge repulsion

between the PNA strand and the DNA strand. Other nucleic acid analogs may contain nucleotides containing polymer backbones, cyclic backbones, or acyclic backbones. For example, the nucleotides may have morpholino backbone structures (U.S. Pat. No. 5,034,506). The analogs may also contain groups such as reporter groups, a group for
5 improving the pharmacokinetic or pharmacodynamic properties of nucleic acid sequence.

In another aspect, the invention provides chimeric nucleic acid sequences comprising a first nucleic acid sequence obtained from flax operatively linked to a second nucleic acid sequence non-native to said first nucleic acid sequence wherein said first nucleic acid sequence comprises a novel flax seed-specific promoter. Preferably, the promoter is
10 selected from the group of promoters comprising Figure 1, Figure 2, Figure 3 and Figure 4 or a nucleic acid sequence hybridizing thereto under stringent conditions.

In accordance with the present invention, the chimeric nucleic acid sequences can be incorporated in a known manner in a recombinant expression vector which ensures good expression in the seed cell. Accordingly, the present invention includes a recombinant
15 expression vector comprising a chimeric nucleic acid sequence of the present invention suitable for expression in a seed cell.

The term "suitable for expression in a seed cell" means that the recombinant expression vectors contain the chimeric nucleic acids sequence of the invention, a regulatory region and a termination region, selected on the basis of the seed cell to be used for
20 expression, which is operatively linked to the nucleic acid sequence encoding the polypeptide of desirable amino acid composition. Operatively linked is intended to mean that the chimeric nucleic acid sequence encoding the polypeptide is linked to a regulatory sequence and termination region which allows expression in the seed cell. A typical construct consists, in the 5' to 3' direction of a regulatory region complete with a promoter
25 capable of directing expression in a plant, a polypeptide coding region and a transcription termination region functional in plant cells. These constructs may be prepared in accordance with methodology well known to those of skill in the art of molecular biology (see for example: Sambrook *et al.* (1990), Molecular Cloning, 2nd ed. Cold Spring Harbor Press). The preparation of constructs may involve techniques such as restriction digestion, ligation, gel
30 electrophoresis, DNA sequencing and PCR. A wide variety of cloning vectors is available to perform the necessary cloning steps. Especially suitable for this purpose are the cloning vectors with a replication system that is functional in *Escherichia coli* such as pBR322, the pUC series M13mp series, pACYC184, pBluescript etc. Nucleic acid sequences may be introduced into these vectors and the vectors may be used to transform *E. coli* which may be
35 grown in an appropriate medium. Plasmids may be recovered from the cells upon harvesting and lysing the cells. Final constructs may be introduced into plant vectors compatible with integration into the plant such as the Ti and Ri plasmids.

The methods for the expression of non-native genes in flax seeds in accordance with the present invention may be practiced using any flax seed-specific promoter and are not limited by the specific flax seed specific promoter that is selected. In preferred embodiments of the present invention, the flax seed-specific promoter confers to the non-native nucleic acid sequence at least one expression characteristic which is similar or identical to an expression characteristic conferred to the native nucleic acid sequence by the native promoter. The term "expression characteristic" as used herein refers to any measurable property or effect conferred by the flax seed-specific promoter to the nucleic acid sequence operably linked to the flax seed-specific promoter. Thus in preferred embodiments, timing of expression in the plant's life cycle, of the non-native nucleic acid sequence is similar or identical to timing of expression of the native nucleic acid sequence. In further preferred embodiments, the expression level of the heterologous nucleic acid sequence is similar or identical to the expression level of the native nucleic acid sequence. In yet further specific embodiments, the response of the non-native gene to alterations in lighting conditions, changes in wavelength or light intensity for example, changes in temperature, tissue wounding, changes in concentration of chemical agents, such as for example phytohormones and pesticides, is similar to the response of the native nucleic acid sequence to these stimuli. Other desired expression characteristics conferred by a flax seed-specific promoter may be recognized by those skilled in the art and a flax seed-specific promoter may be selected accordingly.

Flax-seed specific promoters that may be used in accordance with the present invention include promoters associated with seed storage proteins, such as all albumins and globulins, including the vicilin and legumin-like proteins, as well as seed-specific promoters not associated with seed storage proteins, such as oleosins. Of further particular interest are promoters associated with fatty acid metabolism, such as acyl carrier protein (ACP), saturases, desaturases, elongases and the like.

In preferred embodiments of the present invention the seed specific promoter used is an oleosin promoter, a legumin-like seed storage protein promoter or a 2S storage protein promoter. In particularly preferred embodiments the seed specific promoter has the sequence shown in Figure 1, Figure 2, Figure 3 or Figure 4 or any nucleic acid sequences obtainable from flax and hybridizing to any one of these four nucleic acid sequences under stringent conditions.

Additional flax seed-specific promoters may be used in accordance with the present invention. These promoters may be obtained in a number of ways. Where a flax seed protein has been isolated, it may be partially sequenced, so that a nucleic acid probe may be designed for identifying RNA specific to the seed. To further enhance the RNA specifically associated with the seed, cDNA may be prepared from seed cells and the cDNA may be subtracted with mRNA or cDNA from non-seed cells. The remaining seed

cDNA may then be used to probe a genomic DNA library for complementary sequences. Sequences hybridizing to the cDNA may subsequently be obtained and the associated promoter region may be isolated. It is also possible to screen genomic DNA libraries prepared from flax seed tissues using known seed specific genes from other plant species and subsequently isolate their associated promoters. Due to the relative abundance of seed-storage proteins in seeds it is also possible to obtain sequence information through random sequencing of flax seed cDNA libraries. Those cDNA sequences matching sequence of known seed-storage proteins could be used to identify the associated promoter. Databases containing sequence information from large scale sequencing from for example *Arabidopsis* and maize may be searched for known seed-specific proteins and/or promoters and the information may be used to identify promoter sequences in flax that share sequence similarity. Alternative methods to isolate additional flax seed specific promoters may be used and novel flax seed specific promoters may be discovered by those skilled in the art and used in accordance with the present invention.

The nucleic acid sequence of interest linked to the promoter may be any nucleic acid sequence of interest including any RNA or DNA sequence encoding a peptide or protein of interest, for example, an enzyme, or a sequence complementary to a genomic sequence, where the genomic sequence may be at least one of an open reading frame, an intron, a non-coding leader sequence, or any sequence where the complementary sequence will inhibit transcription, messenger RNA processing, for example splicing or translation. The nucleic acid sequence of interest may be synthetic, naturally derived or a combination thereof. As well, the nucleic acid sequence of interest could be a fragment of the natural sequence, for example just include the catalytic domain or a structure of particular importance. Depending upon the nature of the nucleic acid sequence of interest, it may be desirable to synthesize the sequence with plant preferred codons. The plant preferred codons may be determined from the codons of highest frequency in the proteins expressed in the largest amount in particular plant species of interest.

The nucleic acid sequence of interest may encode any of a variety of recombinant proteins. Examples of recombinant proteins which might be expressed by the methods of the present invention include proteins with a favorable catalytic function or a valuable protein that will accumulate to high levels and be extracted if desired. Proteins with a catalytic function, include, but are not limited to, proteins that confer a new biochemical phenotype on the developing seeds. New phenotypes could include such modifications as altered seed-protein or seed oil composition or seed polysaccharide composition, enhanced production of pre-existing desirable products or properties and the reduction or even suppression on an undesirable gene product using antisense, ribozyme or co-suppression technologies (Izant and Weintraub (1984) Cell 26: 1007-1015, antisense; Hazelhoff and

Gerlach (1988) Nature 334: 585-591, ribozyme; Napoli *et al.* (1990) Plant Cell 2: 279-289, co-suppression).

It is expected that the desired proteins would be expressed in all embryonic tissues, although varying cellular expression may be detected in the different embryonic tissues such as the embryonic axis and cotyledons. The nucleic acid sequence of interest may be expressed at any stage in seed development. The timing of expression may depend on the particular use of the invention. Expression of enzymes involved in oil modification may be desirable early in seed development, for example before accumulation of seed storage protein.

Besides the promoter region and the nucleic acid sequence of interest, a nucleic acid sequence capable of terminating transcription is typically included in expression vectors. Transcriptional terminators are preferably about 200 to about 1,000 nucleotide base pairs and may comprise any such sequences functional in plants, such as the nopaline synthase termination region (Bevan *et al.*, (1983) Nucl. Acid. Res. 11: 369-385), the phaseolin terminator (van der Geest *et al.*, (1994) Plant J. 6(3): 413-423), the terminator for the octopine synthase gene of *Agrobacterium tumefaciens* or other similarly functioning elements. These transcription terminator regions can be obtained as described by An (1987), Methods in Enzym. 153: 292 or are already present in plasmids available from commercial sources such as ClonTech, Palo Alto, California. The choice of the appropriate terminator may have an effect of the rate of transcription.

The chimeric construct may further comprise enhancers such as the AMV leader (Jobling and Gehrke (1987), Nature 325: 622-625) or introns. It should be understood that the design of the expression vector may depend on such factors as the choice of the plant species and/or the type of polypeptide to be expressed.

The expression vectors will normally also contain a marker gene. Marker genes comprise all genes that enable distinction of transformed plant cells from non-transformed cells, including selectable and screenable marker genes. Conveniently, a marker may be a resistance marker to a herbicide, for example, glyphosate or phosphinothricin, or to an antibiotic such as kanamycin, G418, bleomycin, hygromycin, chloramphenicol and the like, which confer a trait that can be selected for by chemical means. Screenable markers may be employed to identify transformants through observation. They include but are not limited to the β -glucuronidase or *uidA* gene, a β -lactamase gene or a green fluorescent protein (Niedz *et al.* (1995) Plant Cell Rep. 14: 403).

In order to introduce nucleic acid sequences into plant cells in general a variety of techniques are available to the skilled artisan. *Agrobacterium*-mediated transformation for flax plant cells has been reported and flax transformants may be obtained in accordance with the methods taught by Dong and McHughen (1993) Plant Science 88: 61-77, although a

variety of other techniques (see below) may also be used to introduce the chimeric DNA constructs in flax cells if so desired.

Transformed flax plants grown in accordance with conventional agricultural practices known to a person skilled in the art are allowed to set seed. Flax seed may then be
5 obtained from mature flax plants and analyzed for desired altered properties with respect to the wild-type seed.

Two or more generations of plants may be grown and either crossed or selfed to allow identification of plants and strains with desired phenotypic characteristics including production of the recombinant polypeptide. It may be desirable to ensure
10 homozygosity in the plants to assure continued inheritance of the recombinant trait. Methods for selecting homozygous plants are well known to those skilled in the art of plant breeding and include recurrent selfing and selection and anther and microspore culture. Homozygous plants may also be obtained by transformation of haploid cells or tissues followed by regeneration of haploid plantlets subsequently converted to diploid plants by
15 any number of known means (e.g. treatment with colchicine or other microtubule disrupting agents).

The present invention also includes transgenic flax seeds prepared according to a method comprising:

- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3'
20 direction of transcription as operably linked components:
- (1) a seed-specific promoter obtained from flax; and
 - (2) a nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said seed-specific promoter;
- (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
25 (c) growing said flax plant cell into a mature flax plant capable of setting seed wherein said nucleic acid sequence of interest is expressed in the seed under the control of said seed-specific promoter.

In preferred embodiments of the invention the seed-specific promoter is selected from the group of flax seed specific promoters consisting of, a 2S storage protein promoter, a
30 globulin promoter, an oleosin promoter, and a legumin-like seed storage protein promoter. Specific promoter sequences are shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) and Figure 4 (SEQ.ID.NO.:8).

The present invention further provides flax plants capable of setting seed prepared by a method comprising:

- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3'
35 direction of transcription as operably linked components:
- (1) a seed-specific promoter obtained from flax; and

- (2) a nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said seed-specific promoter;
- (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
- (c) growing said flax plant cell into a mature flax plant capable of setting seed
- 5 wherein said nucleic acid sequence of interest is expressed in the seed under the control of said seed-specific promoter.

The present invention further provides methods of use for the novel promoters shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) and Figure 4 (SEQ.ID.NO.:8) in plant species other than flax. Accordingly, the invention also

10 includes the preparation of chimeric nucleic acid constructs comprising a promoter selected from the group promoters shown in Figure 1, Figure 2, Figure 3 and Figure 4 and a nucleic acid sequence of interest, and expression in a seed-specific manner of the nucleic acid sequence of interest in plant species other than flax and under the control of the flax promoter.

In another aspect of the present invention there is provided a method for the

15 expression of a nucleic acid sequence of interest in plant seeds comprising:

- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3' direction of transcription as operably linked components;
- (1) a seed-specific promoter selected from the group of seed-specific promoters consisting of
- 20 (i) a nucleic acid sequence as shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8) wherein T can also be U;
- (ii) a nucleic acid sequence that is complimentary to a nucleic acid sequence of (i);
- 25 (iii) a nucleic acid sequence that has substantial sequence homology to a nucleic acid sequence of (i) or (ii); and
- (iv) a nucleic acid sequence that is an analog of a nucleic acid sequence of (i), (ii) or (iii);
- (v) a nucleic acid sequence that hybridizes to a nucleic acid sequence of
- 30 (i), (ii), (iii) or (iv) under stringent hybridization conditions; and
- (2) said nucleic acid of interest;
- (b) introducing the chimeric nucleic acid construct into a plant cell;
- (c) growing said plant cell into a mature plant capable of setting seed, wherein said nucleic acid sequence of interest is expressed in the seed under the control of said
- 35 seed-specific promoter.

A variety of techniques are available for the introduction of nucleic acid sequences, in particular DNA, into plant host cells in general. For example, the chimeric DNA constructs may be introduced into host cells obtained from dicotyledonous plants, such

as tobacco, and oleoagenous species, such as *Brassica napus* using standard *Agrobacterium* vectors by a transformation protocol such as described by Moloney *et al.* (1989), Plant Cell Rep. 8: 238-242 or Hinchee *et al.* (1988) Bio/Technol. 6: 915-922; or other techniques known to those skilled in the art. For example, the use of T-DNA for transformation of plant cells

5 has received extensive study and is amply described in EP 0 120 516, Hoekema *et al.*, (1985), Chapter V In: *The Binary Plant Vector System* Offset-drukkerij Kanters BV, Alblasterdam); Knauf *et al.* (1983), *Genetic Analysis of Host Expression by Agrobacterium*, p. 245, In: *Molecular Genetics of Bacteria-Plant Interaction*, Puhler, A. ed. Springer-Verlag, NY); and An *et al.*, (1985), (EMBO J., 4: 277-284). *Agrobacterium* transformation may also be

10 used to transform monocot plant species (US Patent 5,591,616).

Conveniently, explants may be cultivated with *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* to allow for the transfer of the transcription construct in the plant host cell. Following transformation using *Agrobacterium* the plant cells are dispersed into an appropriate medium for selection, subsequently callus, shoots and eventually plants

15 are recovered. The *Agrobacterium* host will harbour a plasmid comprising the vir genes necessary for transfer of the T-DNA to plant cells. For injection and electroporation (see below) disarmed Ti-plasmids (lacking the tumour genes, particularly the T-DNA region) may be introduced into the plant cell.

The use of non-*Agrobacterium* techniques permits the use of constructs described

20 herein to obtain transformation and expression in a wide variety of monocotyledonous and dicotyledonous plant species. These techniques are especially useful for transformation of plant species that are intractable in an *Agrobacterium* transformation system. Other techniques for gene transfer include particle bombardment (Sanford, (1988), Trends in Biotechn. 6: 299-302), electroporation (Fromm *et al.*, (1985), PNAS USA, 82: 5824-5828;

25 Riggs and Bates, (1986), PNAS USA 83: 5602-5606), PEG mediated DNA uptake (Potrykus *et al.*, (1985), Mol. Gen. Genetics., 199: 169-177), microinjection (Reich *et al.*, Bio/Techn. (1986) 4:1001-1004) and silicone carbide whiskers (Kaeppeler *et al.* (1990) Plant Cell Rep. 9: 415-418).

In a further specific applications such as to *B. napus*, the host cells targeted to

30 receive recombinant DNA constructs typically will be derived from cotyledonary petioles as described by Moloney *et al.* (1989) Plant Cell Rep. 8: 238-242. Other examples using commercial oil seeds include cotyledon transformation in soybean explants (Hinchee *et al.*, (1988) Bio/Technol. 6: 915-922) and stem transformation of cotton (Umbeck *et al.*, (1987) Bio/Technol. 5: 263-266).

Following transformation, the cells, for example as leaf discs, are grown in

35 selective medium. Once the shoots begin to emerge, they are excised and placed onto rooting medium. After sufficient roots have formed, the plants are transferred to soil. Putative transformed plants are then tested for presence of a marker. Southern blotting is performed

on genomic DNA using an appropriate probe, to show integration into the genome of the host cell.

The methods provided by the present invention can be used in conjunction a broad range of plant species. Particularly preferred plant cells employed in accordance with the present invention include cells from the following plants: soybean (*Glycine max*), rapeseed (*Brassica napus*, *Brassica campestris*), sunflower (*Helianthus annuus*), cotton (*Gossypium hirsutum*), corn (*Zea mays*), tobacco (*Nicotiana tabacum*), alfalfa (*Medicago sativa*), wheat (*Triticum* sp.), barley (*Hordeum vulgare*), oats (*Avena sativa* L.), sorghum (*Sorghum bicolor*), *Arabidopsis thaliana*, potato (*Solanum* sp.), flax/linseed (*Linum usitatissimum*), safflower (*Carthamus tinctorius*), oil palm (*Eleais guineensis*), groundnut (*Arachis hypogaea*), Brazil nut (*Bertholletia excelsa*) coconut (*Cocos nucifera*), castor (*Ricinus communis*), coriander (*Coriandrum sativum*), squash (*Cucurbita maxima*), jojoba (*Simmondsia chinensis*) and rice (*Oryza sativa*).

The present invention has a variety of uses which include improving the intrinsic value of plant seeds by their accumulation of altered polypeptides or novel recombinant peptides or by the incorporation or elimination or a metabolic step. Use of the invention may result in improved protein quality (for example, increased concentrations or essential or rare amino acids), improved liquid quality by a modification of fatty acid composition, or improved or elevated carbohydrate composition. Examples include the expression of sulfur-rich proteins, such as those found in lupins or brazil nuts in a seed deficient in sulphurous amino acids. Improved protein quality could also be achieved by the expression of a protein or a fragment of a protein that is enriched in essential amino acids including lysine, cysteine, methionine and tryptophan. Alternatively, a fatty acyl coenzyme A, a transferase enzyme capable of modifying fatty acid ratios in triglycerides (storage lipid) could be expressed. In cases where a recombinant protein is allowed to accumulate in the seed, the protein could also be a peptide which has pharmaceutical or industrial value. In this case the peptide could be extracted from the seed and used in crude or purified form as appropriate for the intended use. As well, the polypeptides that are expressed in the seeds can be fragments or derivatives or the native protein. Pharmaceutically useful proteins may include, but are not limited to, anticoagulants, such as hirudin, antibodies, including monoclonal antibodies and antibody fragments, vaccines, cytokines or growth factors such as bovine growth factor, cholinergic differentiation factor (CDF), ciliary neurotrophic factor (CNTF), fibroblast growth factor (FGF), fish growth factor, gonadotropin, granulocyte-colony stimulating factor (G-CSF), granulocyte-macrophage colony-stimulating factor (GM-CSF), human growth hormone, interferon alpha (IFN- α), interferon beta (IFN- β), interferon gamma (IFN- γ), interleukin 1-alpha (IL1- α), interleukin 1-beta (IL1- β), interleukin-2 (IL-2), interleukin-3 (IL-3), interleukin-4 (IL-4), interleukin-5 (IL-5), interleukin-6 (IL-6), interleukin-10 (IL-10), leukemia inhibitory factor

(LIF), thioredoxin, macrophage colony-stimulating factor (M-CSF), myelomonocytic growth factor, nerve growth factor (NGF), oncostatin M, platelet-derived growth factor (PDGF), prolactin, transforming growth factor alpha (TGF- α), transforming growth factor beta2 (TGF- β 2), tumour necrosis factor alpha (TNF- α), and tumour necrosis factor beta (TNF- β). Pharmaceutically useful proteins can also include mammalian proteins, for example, but not limited to α -1-antitrypsin, anti-obesity proteins, blood proteins, collagen, collagenase, elastin, elastase, enteropeptidase, fibrinogen, haemoglobin, human serum albumin, insulin, lactoferrin, myoglobin and pulmonary surfactant proteins.

Industrially useful peptides may include, but are not limited to α -amylase or other amylases, amyloglucosidase, arabinase, catalase, cellobiohydrolase, cellulases, chitinases, chymotrypsin, dehydrogenases, endo-glucanase, chymosin, endo-galactanase, esterases, β -galactosidase, α -galactosidase or other galactosidases, gastric lipases, glucanases, glucose isomerase, hemi-cellulases, hydrolases, isomerase, ligninases, lipases, lyases, lysozymes, oxidases, oxidoreductase, papain, pectinases, pectin lyase, peroxidases, phosphatases, phytase, proteases, pullulanases, reductases, serine proteases, thioredoxin, transferase, trypsin, and xylanase.

The following non-limiting examples are illustrative of the present invention:

EXAMPLES

EXAMPLE 1

Isolation of Seed-Specific Flax Promoters

Seed specific cDNA clones were isolated from a flax seed specific cDNA-library. These cDNA clones were sequenced and the Basic Local Alignment Search Tool (BLAST) was used to compare these sequences against others in public databases such as Genbank. This comparison revealed that the deduced amino acid sequence of several of the isolated cDNAs had a high degree of similarity to both the low and high molecular weight class of oleosins, 2S-albumin and legumin-like storage proteins. Probes were prepared individually from (portions of) cDNAs encoding oleosins, 2S albumin and legumin-like storage proteins and these were used to screen a genomic library prepared from the flax line Forge that is homozygous for four rust resistance genes (Anderson *et al.* (1997), The Plant Cell 9: 641-651). Several positive lambda clones for each probe were identified after high-stringency screening. The inserts were subcloned into the plasmid vector pBluescript and sequenced. Sequence information revealed that we had isolated the genomic counterparts to the oleosins, 2S albumin and cDNAs *legumin-like* cDNAs. Sequence information of the genomic clones containing sequences encoding a high and low molecular weight oleosin isoforms, 2S albumin and a legumin-like gene are presented in Figures 1 to 4 respectively.

Figure 1 and SEQ.ID.NO.:1 shows the DNA sequence of a flax genomic clone encoding a 16.0 kDa oleosin protein (low molecular weight or L-isoform). Putative regulatory elements are identified and indicated. These include inverted repeats (base

pairs 805 to 813 and 821 to 829; base pairs 1858 to 1866 and 1877 to 1885); direct repeats (base pairs 184 to 193 and 1102 and 1111; base pairs 393 to 402 and 1701 to 1710; base pairs 683 to 692 and 1546 to 1555; base pairs 770 to 781 and 799 to 810; base pairs 955 to 964 and 1936 to 1945; base pairs 1483 to 1496 and 1513 to 1526), the abscisic acid responsive element (ABRE) (base pairs 1859 to 1866), CACA box (base pairs 1933 to 1936), TATA box (base pairs 1925 to 1931) and CAT box (base pairs 1989 to 1993). As well, the poly adenylation signal is indicated (base pairs 3020 to 3025). The open reading frame is interrupted by 1 short intron (which are marked) and the 2 exons are translated and indicated in IUPAC single letter amino-acid codes.

Figure 2 and SEQ.ID.NO.:4 shows the DNA sequence of a flax genomic clone encoding a 18.6 kDa oleosin protein (high molecular weight or H-isoform). Putative regulatory elements are identified and indicated. These include direct repeats (base pairs 14 to 25 and 1427 to 1438; base pairs 80 to 89 and 1242 to 1251; base pairs 177 to 186 and 837 to 846; base pairs 1281 to 1290 and 1242 to 1251; base pairs 1591 to 1600 and 1678 to 1287). The open reading frame is not interrupted by introns and is translated and indicated in IUPAC single letter amino-acid codes.

Figure 3 and SEQ.ID.NO.:6 shows the DNA sequence of the flax genomic clone encoding a 2S storage protein. Nucleotide sequencing of this clone revealed it to have an open reading frame of 174 amino acids that showed homology to the plant 2S storage group of proteins. The sequence encodes an open reading frame with 38% overall similarity to a *Brassica oleracea* 2S storage protein, including complete conservation of the glutamine-rich stretch QQQGQQGQQQ (SEQ.ID.NO.:13). In addition, the 2S storage protein gene promoter contained several putative promoter regulatory elements. These include AT rich repeats (base pairs 25-36, 97-108 and 167-190), RY-like repeat (base pairs 240-247), G-box-like element (base pairs 274-280), seed specific box-like motif (base pairs 285-290) and TATA box (base pairs 327-333).

Figure 4 and SEQ.ID.NO.:8 shows the DNA sequence of a flax genomic clone encoding a 54.4 kDa flax legumin-like seed storage protein. This legumin-like seed storage protein gene will also be referred to as "linin". The deduced amino acid sequence of the linin gene was compared to the *legumin-like* protein from *R. communis*, the legumin precursor from *M. salicifolia*, *Q. robur* and *G. hirsutum*, the glutelin precursor from *O. sativa* and a 12 S seed storage protein from *A. thaliana*. The linin gene shows a sequence identity/similarity with the corresponding proteins from *R. communis*, *M. salicifolia*, *Q. robur*, *G. hirsutum*, *O. sativa* and *A. thaliana* of 59/15, 47/16, 50/17, 45/17, 43/18 and 43/18 percent respectively. Putative regulatory elements in the promoter region are identified and indicated. These include inverted repeats (base pairs 265 to 276 and 281 to 292; base pairs 513 to 524 and 535 to 545), repeats (base pairs 1349 to 1360 and 1367 to 1378; base pairs 1513 to 1529 and 1554 to 1572), the abscisic acid responsive element (ABRE) (base pairs 1223 and

1231), legumin box (RY repeats) (between base pairs 1223 and 1231), a possible vicilin box region (base pairs 1887 to 1894), CAAT box (base pairs 1782 to 1785) and TATA box (base pairs 1966 to 1970). As well, the signal peptide for ER membrane targeting is indicated (base pairs 2034-2080). The open reading frame is interrupted by 3 short introns (which are marked) and the 4 exons are translated and indicated in IUPAC single letter amino-acid codes.

Figure 5 shows Southern blot analysis of flax genomic DNA. 60 µg of flax genomic DNA was isolated from leaves, digested with EcoRI (lane 1), HindIII (lane 2) and BamHI (lane 3) and was loaded into the respective lanes. A) Hybridizations were performed with random primed ³²P-labelled 3T cDNA (high molecular weight flax oleosin isoform). B) Hybridizations were performed with random primed ³²P-labelled 7R cDNA (low molecular weight flax oleosin isoform). The results demonstrate that both 3T (high molecular weight oleosin isoform) and 7R (low molecular weight oleosin isoform) oleosin cDNAs hybridize with flax genomic DNA. More specifically the results indicate that 3T is likely to represent a 2-copy gene in flax, as seen by two bands in each lane of digestion. Similarly, 7R is likely to represent a multigene family in flax as multiple bands were detected for each digestion.

EXAMPLE 2

Seed specific expression of flax oleosin genes

Figure 6 shows a Northern blot analysis of the seed specific expression of flax oleosins. Northern hybridization of the two oleosin mRNA in different tissues. Ten µg of total RNA was extracted from different tissues, R, root; C, cotyledon; L, leaf; S, seed capsule; E, embryo. The membrane was probed with (A) cDNA encoding high molecular weight (H)-isoform (identical to coding sequence as presented in Figure 2) and (B) cDNA encoding low molecular weight (L) -isoform (identical to coding sequence as presented in Figure 1). Both the transcripts are expressed only in the embryo and seed capsule, which contains embryos.

EXAMPLE 3

Developmental expression of flax oleosin genes during seed development

Figure 7 shows a Northern blot analysis of the developmental expression of flax oleosins during seed development. 15 µg per lane of total RNA was loaded in each lane on agarose/formaldehyde gel and blotted onto HybondN+ membrane. 10j: This membrane was probed using the ³²P dCTP labeled flax oleosin cDNA clone (low molecular weight isoform). Stages indicated are the number of days past anthesis (DPA). 3T) 15 µg per lane of total RNA was loaded in each lane on agarose/formaldehyde gel and blotted onto HybondN+ membrane. 3T: This membrane was probed using the ³²P dCTP labeled flax oleosin cDNA clone (high molecular weight isoform). Both the transcripts were expressed very early in

development (6DPA, early cotyledonary stage). Expression is maximum at 16 to 20 DPA (late cotyledonary stage) and declines at 22 DPA (mature embryos).

EXAMPLE 4

Transient Seed specific expression of β -glucuronidase (GUS) when under the regulatory control of flax oleosin regulatory sequences

Two constructs were made using standard molecular biology techniques (eg see Sambrook *et al.* (1990), Molecular Cloning, 2nd ed. Cold Spring Harbor Press, including restriction enzyme digestions, ligation and polymerase chain reaction (PCR).

Construct pSC54: The β -glucuronidase reporter coding sequence from vector GUSN358>S (Clontech Laboratories) was placed between the promoter sequence from nucleotide 21 to 1852 and terminator sequence from 2395 to 3501 (as described in Figure 1). This insert was cloned into pBluescript and the resulting vector is called pSC54

Construct pSC60: The β -glucuronidase reporter coding sequence from vector GUSN358>S (Clontech Laboratories) was placed between the promoter sequence from nucleotide 1 to 2023 and terminator sequence from 2867 to 3925 (as described in Figure 2). This insert was cloned into pBluescript and the resulting vector is called pSC60.

pSC54, pSC60 and a promoter-less GUS construct (Control) were introduced into the flax embryos using particle bombardment using standard protocols (eg see Abenes *et al.* (1997) Plant Cell reports 17:1-7). Figure 8 shows the GUS activity of flax embryos bombarded with pSC54, pSC60 and a promoterless GUS construct measured 48 hours after particle bombardment. As can be seen the flax oleosin regulatory sequences are sufficient to drive the expression of GUS in flax embryos.

EXAMPLE 5

Stable seed specific expression of β -glucuronidase (GUS) in flax and Arabidopsis when under the regulatory control of flax 2S storage protein gene promoter

A GUS reporter gene construct was made by incorporating 5' and 3' regions from the DNA fragment described in Figure 3 into promoterless-GUS pBI101 vector as follows.

A 400bp amplicon from the 5' end of the DNA fragment described in Figure 3 was PCR amplified using the following primers (location shown in Fig 3):

5' primer(1): 5'-TCCAATATGTAGGTCATA-3' (SEQ.ID.NO.:14)

3' primer(1): 5'-CTTTAAGGTGTGAGAGTC-3' (SEQ.ID.NO.:15)

The PCR primers also contained restriction sites for HindIII and BamHI which were used to clone the 400bp 5'UTR amplicon into the HindIII/BamHI sites of the pBI101 vector in front of the GUS reporter gene. A 736bp amplicon from the 3' untranslated region (3'UTR) of the DNA fragment described in Figure 3 was PCR amplified using the following primers (location shown in Fig 3):

5' primer (2): 5'-AGGGGTGATCGATTA-3' (SEQ.ID.NO.:16)

3' primer (2): 5'-GATAGAACCCACACGAGC-3' (SEQ.ID.NO.:17)

The PCR primers also contained restriction sites for *SacI* and *EcoRI*. The NOS terminator region of the pBI101 vector was cut out with *SacI*/*EcoRI* digestion and replaced with the similarly digested 736bp 3'UTR amplicon of the DNA fragment described in Figure 3.

5 The GUS reporter construct was then electroporated into *Agrobacterium tumifaciens* strain AGL1 and transformation of flax (Finnegan *et al.* (1993) Plant Mol Biol. 22(4): 625-633) and *Arabidopsis* (Valvekens *et al.* Proc. Natl. Acad. Sci. 85: 5536-5540) carried out according to previously described protocols.

10 Various tissues from flax and *Arabidopsis* plants carrying the GUS reporter construct were assayed histologically for evidence of GUS activity. In the case of flax, leaf tissue, root tissue and mid-maturity embryos dissected out of developing seeds were stained for GUS activity. For *Arabidopsis*, developing seeds were stained for GUS *in situ* in their siliques.

15 GUS staining was carried out by immersing the tissues in histochemical buffer containing 0.5 mM X-gluc, 0.5 M potassium phosphate buffer (pH 7.0), 1 mM EDTA, 0.5 M sorbital, 0.5 mM potassium ferricyanide and 0.5 mM potassium ferrocyanide. The staining reaction was carried out for 12-16 hrs at 37°C and the reaction was stopped by adding 95% ethanol. Tissues were subsequently cleared of chlorophyll by repeated washing in 95% ethanol prior to photography. Figure 9 shows clear evidence of strong GUS activity in
20 developing flax embryos and *Arabidopsis* seeds, and no evidence of GUS reporter gene expression in flax roots or leaves, or in *Arabidopsis* silique walls.

EXAMPLE 6

Stable seed specific expression of β -glucuronidase (GUS) in flax, *Arabidopsis* and *Brassica napus* when under the regulatory control of flax legumin-like storage protein gene regulatory sequences
25

A construct was made using standard molecular biology techniques, including restriction enzyme digestions, ligation and polymerase chain reaction (PCR). In order to obtain a DNA fragment containing approximately 2 kilobases from the 5' transcriptional initiation region of the flax legumin-like seed storage protein in a configuration suitable for
30 ligation to a GUS coding sequence, a PCR based approach was used. This involved the use of the polymerase chain reaction to amplify the precise sequence desired for the expression analysis. To perform the necessary PCR amplification, two oligonucleotide primers were synthesized (Beckman Oligo 1000M DNA synthesizer) have the following sequences:

5' primer: 5'**TA**CT**AGACTCAAGCATACGGACAAGGGT** 3' (SJ-634) (SEQ.ID.NO.:18)

35 The italicized bases correspond to nucleotide positions 1 to 21 in the sequence reported in Figure 4. The additional nucleotides 5' of this sequence in the primer are not identical to the promoter sequence, but were included in order to place a *XbaI* site at the 5'

end of the amplification product. The XbaI (5'-TCTAGA-3') (SEQ.ID.NO.:19) site is underlined.

A second (3') primer was synthesized which had the following sequence:

3' primer 5'GGTTATCATTGTATGAACTGA3' (SJ-618) (SEQ.ID.NO.:20)

5 This primer contains the precise complement (shown in italics) to the sequence reported in Figure 4 from bases 2343 to 2363. This primer was not designed with an additional restriction enzyme site due to the fact that a natural NcoI site (5'-CCATGG-3') (SEQ.ID.NO.:21) straddles the start codon between base pairs 2034 and 2039, thereby allowing for insertion of the storage protein promoter into the appropriate cloning vector.

10 These two primers were used in a PCR amplification reaction to produce a DNA fragment containing the sequence between nucleotides 1 and 2342 of the flax seed storage protein gene with a XbaI site at the 5' end and a NcoI site 302 base pairs from the 3' end. PCR amplification was performed using the enzyme Pfu (Stratagene) using conditions recommended by the enzyme manufacturer and a temperature program of 94°C
15 (denaturation) for 1 minute, 55°C (annealing) for 1 minute, and 72°C (elongation) for 3.5 minutes. The template was the legumin seed storage protein genomic clone shown in Figure 4.

The resulting amplification product was subsequently digested with XbaI and NcoI to remove the desired 2 kb promoter region. This promoter fragment was cloned into
20 the XbaI and NcoI sites of a XbaI and NcoI digested plasmid designated pGUS1318 (Plasmid pGUSN358S (Clontech Laboratories) was cut with NcoI and EcoRI and the GUS insert was cloned into pBluescriptKS+ (Stratagene) which was adapted to contain an NcoI site in the multiple cloning site.) The resulting plasmid containing the promoter-GUS fusion was called pPGUS1318. The terminator of the legumin seed storage protein from flax
25 was also amplified from the above mentioned genomic clone. To perform the necessary PCR amplification, oligonucleotide primers were synthesized having the following sequences:

5' primer: 5' GCAAGCTTAATGTGACGGTGAAATAATAACGG 3' (SJ620) (SEQ.ID.NO.:22)

The italicized bases correspond to nucleotide positions 3780 to 3803 in the sequence reported in Figure 4. The additional nucleotides 5' of this sequence in the primer are not
30 identical to the promoter sequence, but were included in order to place a HindIII site at the 5' end of the amplification product. The HindIII site (5'-AAGCTT-3') (SEQ.ID.NO.:23) is underlined.

A second (3') primer was synthesized which had the following sequence:

3' primer 5'TAGGTACCTGGCAGGTAAAGACTCTGCTC3' (SJ-618) (SEQ.ID.NO.:24)

35 This primer contains the precise complement (shown in italics) to the sequence reported in Figure 4 from bases 4311 to 4290. The additional nucleotides 5' of this sequence in the primer are not identical to the promoter sequence, but were included in order to place a

KpnI site at the 5' end of the amplification product. The KpnI site (5'-GGTACC-3') (SEQ.ID.NO.:25) is underlined.

These two primers were used in a PCR amplification reaction to produce a DNA fragment containing the sequence between nucleotides 3779 and 4311 of the flax seed storage protein gene terminator with a HindIII site at the 5' end and a KpnI site at 3' end. Amplification using PCT was as described above. The above pPGUS1318 vector that contains the amplified promoter was digested with XhoI and treated with Klenow to create a blunt end. The vector was subsequently digested with KpnI and the above amplified terminator sequence was inserted so that it was located 3' of the GUS coding sequence. The resulting vector containing the flax seed storage protein promoter, GUS and the flax seed storage protein terminator is referred to as pPGUST.

The XbaI-KpnI insert of pPGUST which contains the linin promoter-GUS coding sequence-linin terminator sequence was ligated into the XbaI-KpnI sites of pSBS3000 (This vector is a derivative from the *Agrobacterium* binary plasmid pPZP221 (Hajdukiewicz et al., 1994, Plant Molec. Biol. 25: 989-994). In pSBS3000 the plant gentamycin resistance gene of pPZP221 was replaced with parsley ubiquitin promoter-phosphinothricin acetyl transferase gene-parsley ubiquitin termination sequence to confer resistance to the herbicide glufosinate ammonium). The resulting vector is called pSBS2089. In addition the XbaI-KpnI insert of pPGUST which contains the linin promoter-GUS coding sequence-linin terminator sequence was ligated into the XbaI-KpnI sites of the *Agrobacterium* binary plasmid pCGN1559 (MacBride and Summerfield, 1990, Plant Molec. Biol. 14 269-276, confers resistance to the antibiotic kanamycin)). The resulting vector was called pSBS2083. Plasmids pSBS2089 and pSBS2083 were electroporated into *Agrobacterium* strain EHA101. *Agrobacterium* strain EHA101 (pSBS2089) was used to transform flax and Arabidopsis, *Agrobacterium* strain EHA101 (pSBS2083) was used to transform *Brassica napus*. Flax transformation was performed essentially as described in Jordan and McHuguen (1988) Plant cell reports 7: 281-284, except transgenic shoots were selected on 10 µM L-phosphinothricine instead of kanamycin. Arabidopsis transformation was done essentially as described in "Arabidopsis Protocols; Methods in Molecular Biology" Vol 82. Edited by Martinez-Zapater JM and Salinas J. ISBN 0-89603-391-0 pg 259-266 (1998) except the putative transgenic plants were selected on agarose plates containing 80µM L-phosphinothricine. *Brassica napus* transformation was done essentially as described in Moloney et al. (1989). Plant Cell Reports. 8: 238-242.

Figure 10 shows the tissue-specific expression of GUS in transgenic flax plants transformed with a linin-GUS gene construct (pSBS2089). GUS expression was measured in roots (R), stems (S), leaves (L), Buds (B) and embryo (E). Some expression was seen in buds, and maximal expression was achieved in embryo tissues. No detectable expression was seen in any of the untransformed (WT) tissues.

Figure 11 shows the temporal expression of GUS in transgenic flax plants transformed with a linin-GUS gene construct (pSBS2089). As can be seen, maximum expression is achieved in mature (pre-dessicated) flax embryos.

5 Figure 12 shows the absolute expression of GUS in transgenic *Brassica napus* plants (L1 to L9) transformed with a linin-GUS gene construct (pSBS2083). As can be seen high level expression can be achieved in *Brassica napus* plants. When comparing individual transgenic plants, a typical variation in expression due to position effect can also be seen.

10 Figure 13 shows expression of GUS in transgenic *Arabidopsis* siliques (transformed with a linin-GUS gene construct (pSBS2089)) during seed development. As can be seen high level expression can also be achieved in *Arabidopsis* seed tissues. Maximum expression is achieved at stage 4 (mature but not fully dessicated) of seed development. No detectable expression is observed in non-seed tissues such as leaves, stems, roots and silique walls (results not shown).

15 While the present invention has been described with reference to what are presently considered to be the preferred examples, it is to be understood that the invention is not limited to the disclosed examples. To the contrary, the invention is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims.

20 All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

We Claim:

1. A method for the expression of a nucleic acid sequence of interest in flax seeds comprising:
 - (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3' direction of transcription as operably linked components:
 - (1) a seed-specific promoter obtained from flax; and
 - (2) said nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said seed-specific promoter;
 - (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
 - (c) growing said flax plant cell into a mature flax plant capable of setting seed wherein said nucleic acid sequence of interest is expressed in the seed under the control of said seed-specific promoter.
2. The method according to claim 1 wherein at least one expression characteristic conferred by said seed-specific promoter to its native nucleic acid sequence is conferred to said non-native nucleic acid sequence.
3. The method according to claim 2 wherein said expression characteristic is timing of expression, level of expression, response to a change in lighting conditions, response to a change in temperature, response to a change in concentration of a chemical agent.
4. The method according to claim 1 wherein said flax seed-specific promoter is selected from the group of promoters comprising, oleosin promoters, 2S storage protein promoters and legumin-like seed-storage protein promoters.
5. The method according to claim 1 wherein said flax seed-specific promoter comprises:
 - (a) a nucleic acid sequence as shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8) wherein T can also be U;
 - (b) a nucleic acid sequence that is complimentary to a nucleic acid sequence of (a);
 - (c) a nucleic acid sequence that has substantial sequence homology to a nucleic acid sequence of (a) or (b);
 - (d) a nucleic acid sequence that is an analog of a nucleic acid sequence of (a), (b) or (c); or
 - (e) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (a), (b), (c) or (d) under stringent hybridization conditions.

6. The method according to claim 1 wherein expression of said nucleic acid sequence of interest results in an alteration in protein or fatty acid composition in said seed.
7. Transgenic flax seed prepared according to a method comprising:
- 5 (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3' direction of transcription as operably linked components:
- (1) a seed-specific promoter obtained from flax; and
- (2) a nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said seed-specific promoter;
- (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
- 10 (c) growing said flax plant cell into a mature flax plant capable of setting seed wherein said nucleic acid sequence of interest is expressed in the seed under the control of said seed-specific promoter.
8. Transgenic flax seed according to claim 7 wherein at least one expression characteristic conferred by said seed-specific promoter to its native nucleic acid sequence is
- 15 conferred to said non-native nucleic acid sequence.
9. The method according to claim 8 wherein said expression characteristic is timing of expression or level of expression.
10. Transgenic flax seed according to claim 8 wherein said seed-specific promoter is a seed storage protein promoter, an oleosin promoter, a 2S storage protein promoter or a
- 20 legumin-like seed-storage protein promoter.
11. Transgenic flax seed according to claim 8 wherein said seed specific promoter comprises:
- (a) a nucleic acid sequence as shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8)
- 25 wherein T can also be U;
- (b) a nucleic acid sequence that is complimentary to a nucleic acid sequence of (a);
- (c) a nucleic acid sequence that has substantial sequence homology to a nucleic acid sequence of (a) or (b);
- 30 (d) a nucleic acid sequence that is an analog of a nucleic acid sequence of (a), (b) or (c); or

- (e) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (a), (b), (c) or (d) under stringent hybridization conditions.
12. Transgenic flax seed according to claim 8 wherein expression of said non-native gene of interest results in an alteration in the seed protein or fatty acid composition.
- 5 13. Transgenic flax plants capable of setting seed prepared by a method comprising:
- (a) preparing a chimeric nucleic acid construct comprising in the 5' to 3' direction of transcription as operably linked components:
- 10 (1) a seed-specific promoter obtained from flax; and
- (2) a nucleic acid sequence of interest wherein said nucleic acid of interest is non-native to said seed-specific promoter;
- (b) introducing said chimeric nucleic acid construct into a flax plant cell; and
- (c) growing said flax plant cell into a mature flax plant capable of setting seed wherein said nucleic acid sequence of interest is expressed in the seed under the control of
- 15 said seed-specific promoter.
14. An isolated nucleic acid sequence capable of directing seed-specific expression in a plant comprising:
- (a) a nucleic acid sequence as shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8) wherein T can also be U;
- 20 (b) a nucleic acid sequence that is complementary to a nucleic acid sequence of (a);
- (c) a nucleic acid sequence that has substantial sequence homology to a nucleic acid sequence of (a) or (b); or
- 25 (d) a nucleic acid sequence that is an analog of a nucleic acid sequence of (a), (b) or (c); or
- (e) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (a), (b), (c) or (d) under stringent hybridization conditions.
15. An isolated chimeric nucleic acid sequence comprising:
- 30 (a) a first nucleic acid sequence comprising a seed-specific promoter obtained from flax which comprises:
- (1) a nucleic acid sequence as shown in Figure 1 (SEQ.ID.NO.:1), Figure 2 (SEQ.ID.NO.:4), Figure 3 (SEQ.ID.NO.:6) or Figure 4 (SEQ.ID.NO.:8) wherein T can also be U;

- (2) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (a) under stringent hybridization conditions;
- (3) a nucleic acid sequence that is complimentary to a nucleic acid sequence of (a); or
- 5 (4) a nucleic acid sequence that has substantial sequence homology to a nucleic acid sequence of (a); and
- (b) a second nucleic acid sequence non-native to said flax seed-specific promoter.

16. A method for the expression of a nucleic acid sequence of interest in a plant seed comprising:

- (a) introducing the chimeric nucleic acid sequence according to claim 15 into a plant cell; and
- (b) growing said plant cell into a mature plant capable of setting seed,
- wherein the second nucleic acid sequence is expressed in the seed under the control of the
- 15 seed specific promoter.

17. A method according to claim 16 wherein said plant cell is selected from the group of plants consisting of soybean (*Glycine max*), rapeseed (*Brassica napus*, *Brassica campestris*), sunflower (*Helianthus annuus*), cotton (*Gossypium hirsutum*), corn (*Zea mays*), tobacco (*Nicotiana tabacum*), alfalfa (*Medicago sativa*), wheat (*Triticum sp.*), barley

20 (*Hordeum vulgare*), oats (*Avena sativa* L.), sorghum (*Sorghum bicolor*), *Arabidopsis thaliana*, potato (*Solanum sp.*), flax/linseed (*Linum usitatissimum*), safflower (*Carthamus tinctorius*), oil palm (*Eleais guineensis*), groundnut (*Arachis hypogaea*), Brazil nut (*Bertholletia excelsa*) coconut (*Cocus nucifera*), castor (*Ricinus communis*), coriander (*Coriandrum sativum*), squash (*Cucurbita maxima*), jojoba (*Simmondsia chinensis*) and rice

25 (*Oryza sativa*).

18. A plant prepared according to the method of claim 16.

19. A plant cell comprising the chimeric nucleic acid sequence according to claim 15.

20. Plant seed comprising the chimeric nucleic acid sequence according to claim 15.

21. Plant seed obtained from a plant prepared according to the method of claim 16.

30 22. A recombinant expression vector comprising a nucleic acid sequence according to claim 14.

23. A recombinant expression vector comprising a nucleic acid sequence according to claim 15.

FIGURE 1-1

1 ttcaaaacccgattcccagagcgccctattgaagatatgggggaagtccgacgagatcgatgctgggtcagtgctatg 80
 81 gtgatggtgccgttttggggggaggatgagcgagatagccaagactagcattccgttccccacacagagttgggaattgtg 160
 161 ccaaatccaacacttgtcgtattgagcgacgtagggacgcggaaaaacacatccgttgatcagggagttgtacgatg 240
 241 atctcgagccttatgtgtcgaagaatccgagggtatgcttacgtgaactacagggatctcgacatcgggatgaatggagga 320
 321 ggtgaagggtgagaaagggtacttatggtgaggctaaagggtgtgggggagaagtacttttgggggtcaactttgatcgggtt 400
 401 ggttcgggtgaagacgattgttgatcccaataatgtgtttcgaaacgagcagagcattccctcaattccaactcgggttat 480
 481 aaggatcaatgatcaatgagaattttcctttccaatgtgattacaagtctattgggtcagcttttctcaactgctcctat 560
 561 tcattttagatttaataacaactattaattaccagccttttatccggccgttgccgatttattttcttaagtttt 640
 641 agatgaaatgaaaccgatttagttttttattgagatgagattaaatccttaatttgccttgaaatttactcaagggtgatgtga 720
 721 tatttggaaattaactaaaatgataaaatcgataaaaaataaaatattt^{R3}aaaaataaaataaaacataaagaacaata 800
 801 aaataaaataaaattt^{I1}aaattt^{R4}attt^{R5}cttctgtatcatacatctcttcttacttctttaaaggctt 880
 881 ttcaattatcaccttaataataacataagataaaatcgtaaattctataaacctatacacattgcacggtgaacaat 960
 961 caatatgataataataataataataattcaattattaatctacaattttttaattataaaagtttatgcggtcagtt 1040
 1041 tctgcaagctccgagctccttgtcatcgtttagtttctgcggtctcaagggtataacgactcggagcagagcccttgc 1120
 1121 tccaatggacgggttgcatcttctgccgtcgttgagctcgattggcggtgcatgctggagtcagagttcctacaaaaaac 1200
 1201 cctaaactagaggggtgattaggtgaaattaggggttgccctgggtccattgtccaaagtttagtcaacttaaaac 1280
 1281 agacttaaaatttatgcttcaaaaatagttttatctgtttattatattagcgtgtaattagctcttgacaatggggccggacgg 1360

2/24

FIGURE 1-2

1361 gtacggattcgggaccccgatccccgcccatagtgtaatggctcaactgccaagtgcagcattggaccgaaattattggac 1440
 1441 acgaagtactaatgtgaaaaacttttacatttgttattttctactttaatactatgctattttcaaaaatttgaactttaat 1520
 1521 actatgtttttatagtttagtataatcttaatttttatgcaaaattcatctaaattgtattataaactatttttcgatccgtag 1600
 1601 ctaattatttcgaaggcaagtcaaaagtgttattgtggactatgtgagctaatttgaaacctttatctctcccaaccactc 1680
 1681 aagttaattgaaccaaactcgatcggttgggttcgagctatttcgagccattgtgttatatgcacgtgagatatcaag 1760
 1761 attgacccgaacacttttattatgataatgtagaaaaagaaaaacataatttctaagactacatgcatgcaaaagtgcacccct 1840
 1841 gcatggaaagctgctcaaacacggtggcatagactcccgccacggtgtccattccacctcatcacctcaccaccccgttcac 1920
 1921 ctcttattatcacacaatcaatcaatcctactcctcactcctcactcgaacaaatccgaccaacttataccaatattccca 2000
 2001 aacttgattaaatttctcagcaat ATG GAT CAG ACG CAC CAG ACA TAC GCC GGA ACC ACG CAG AAC 2065
 1 M D Q T H Q T Y A G T T Q N 14
 2066 CCG AGC TAT GGC GGC GGC ACA ATG TAC CAG CAG CAG CCG AGG TCT TAC CAG GCG 2125
 15 P S Y G G G T M Y Q Q Q Q P R S Y Q A 34
 2126 GTG AAG GCG GCC ACT GCA GCC ACC GCG GGT GGA TCC CTC ATC GTT CTG TCC GGT CTC ATC 2185
 35 V K A A T A A T A G G S L I V L S G L I 54
 2186 CTT ACG GCC ACC GTC ATT TCA CTC ATC ATA GCC ACC CCT CTC CTC L L V I F S P 2245
 55 L T A T V I S L I I A T P L L V I F S P 74
 2246 GTT CTT GTC CCG GCT CTC ATC ACC GTC GGC CTC TTG ATC ACC GGG TTT CTT GCT TCC GGT 2305
 75 V L V P A L I T V G L L I T G F L A S G 94
 2306 GGG TTC GGA GTC GCC ACC GTC TTG TCC TGG ATC TAT AG gtatgtataagcctttggactt 2370
 95 G F G V A A V T V L S W I Y R 109
 2371 tagtattgttataaaataacataaagctgatttatgaacatggatctctcccaacaagagttatttaaatgcatctcgtctg 2450

3/24

FIGURE 1-3

2451 actcgatcgggttgggttttgagctactcgggtcacaaatggtcgggtcggtctggtatctgttatactaataatttggaaagcc 2530
 2531 tgaagtttcatattgttctgccccaaacttcccactaccttttgagggtgtttaagaagccatacaaaactaattatgaatccct 2610
 2611 cccaacaactcagaactcgagtcagtggtgtgtgacggttctctataaaacatttcgaaaaatcttgttcaatgaacgtag 2690
 2691 aaatgaccatgcttgatgattgtgggtcttataag G TAC GTG ACC GGC GGG CAC CCG GCG GGA GGG 2756
 110 Y V T G G H P A G G 119
 2757 GAT TCG CTG GAC CAG GCT AGG TCG AAG CTG GCC GGA AAG GCC AGG GAG GTG AAG GAC AGG 2816
 120 D S L D Q A R S K L A G K A R E V K D R 139
 2817 GCG TCG GAG TTC GCA CAG CAG CAT GTC ACA GGT GGT CAA CAG ACC TCT TAA agagagtcctct 2879
 140 A S E F A Q Q H V T G G Q Q T S * 156
 2880 agttaaatgttcttctgtttctgttctgtggtggtgtgtaaaactctcttttaagtgtgctgttttcttctgtctgtgt 2959
 2960 gttgtaagtgaagtgtaaatcgaagttccaagttgagatgtttgttaacgatgatgttttcttaataatcagagatattaa 3039
 3040 aagggttgctaatttagtattgctgtctgatctcggaccacaaactcgcaagtaaaattgcagaggatgagttgtacagaaca 3119
 Poly A signal
 3120 agcgtgcattgttctggaagttcatctccttggagccgaccttgttgcagtttccgccaagtcactagacaaatgtt 3199
 3200 acgagtaagcctctgtcaaacagatcgctctagcgtcccgagaaacaccagatttttgcgaaaaccatcggggatcaatt 3279
 3280 ttcgattcaattccgatcttgggaagtacttgaacagaagcatgatgctaaaaagataatagaaaaatcgaaagcctagaaaaag 3359
 3360 ttgtacagaaagcaaaagtcaaaaaatatagatcaacttcaaaggttcaaattacatcttacagacccccaaaaatgaca 3439
 3440 gttaacagaagtcgactaaacagaaaccagccagcttcacctggaatgaaggagctttgatcaatccatccctagcttcat 3519
 3520 tccccctttgaaattgcagacagagctctcatcctgctaaaagctggtggttattcttaacccctgcaatcaataagcatga 3599
 3600 actaacattggacaccttcatcggcggttgcctcgaaaaatcagtgagcgggatttacctgtgtgtgtagtaaacctctc 3679

FIGURE 1-4

3680 tccttgataaaaatctgaaattccggcatcaactactgcccacctttctgcttaaggtgattttatcaccaaggctga 3759
3760 gcgtgattccttgcgcttctgctccgaatcctgatgtatccactgagctttccatctccttcccttctccaggccttatgttc 3839
3840 accaatgcgtcctcgcgcgaacacactcttggcgtacaagtgcagccaggaatccacactctccatcaagtgcagacct 3919
3920 gcaaaccccaataagaacacacaaactccaaagtcacacgatcaattctcgccttttatgaagaaaaaggaaacttctgggt 3999
4000 acttacggtgccgtcagacacttccatattttagacttgatgatatggtccaggaaattccttctcgttctgaattgtgt 4079
4080 gttaacagcaaacctgacagacagaaaagatatcgcaaaatttaagatactgggatgactaggcacagagaaaaatgaaatcttaa 4159
4160 ttctagaagtaaaaccttattttcccatccttctgcccacatagtcgggaacgcagcatccgaggaagaagcaggag 4239
4240 agatgtaatccatgatgcgtgtggatatcgttgaggacgacaaactgaacgttcccatcacattgg 4305

1 tctagacatttgacataaaaccgaattcaaaagaacacaaacttgactaacacccaaagaaatagagttagtgaatttggg_{R1} 80
81 agattaaaaaatagaacaaactgattctcttagaaagaagagatgattagtgctttcagttcgtctgtcaggaatcga_{R2} 160
161 gatgttcacttatttacattgtcgattcatctcccaattgtcctggttccttactgtccgacgcttttttgaatcccag 240
241 ttaattcccatcaagtcttccttcagctgcgtagcactgctagctccaacatggagcgtggagtctactcgttcatgggg 320
321 catcgcaaagggttgcccttcattgttctgctaccagccagcgcacccctcttggtgtgtggacaaattgcggtgaagc 400
401 gcgcaagttgacatcccatagttctgcacacttcaccatatggatgttttaaaacgtatatcacgagtgcgatctacatgtc 480
481 ccatcacaccacatataaagcaatagtttgggagcttttccatatttgaaaacggcattgacgacttgcctctcgcataat 560
561 ttaatctttttctcttcagctgattgtgtgcatccattcgggctcagaagcacatcaaagggatctctccatcgtagt 640
641 attgggtcgtgtcgtatgatacgaagcagtcgatgaagtttccctaattgtgcgagctacaggctccgcaaaagaacccgcga 720
721 ggtagatcgtatgtagtaccacaaaatcagtttgcgtagcggaaatcaacactagagactcacccctaataatgcattctcatg 800
801 tgtgatgaacagtttatcatttgtgagtctaggggtcattgtcgatgaccccaatgcacattgagcttatgatagaatttg_{R3} 880
881 aataggaagcgttttccaccagatcacgaatagctaccccttttttcgggcgcaaatcttcggcatcctatcttccacc 960
961 acaacttaagatcgatcggtaaggaaactcacgcgaccacacatcgaaataatcttcggtgacgggttcctgttgatca 1040
1041 agtccctcaatttcctcaacctagttcttcaatcgccgctagcgttatccccgcataatggactttcatagcgggagcgt 1120
1121 agccggagacgacgagcaagaaggatgagcggcgagattgcgggctaagaacgagcttccctgcttgccttatggag 1200
1201 gcgatttctgagttgatgggtgatgattgtgtgatgtggacacttttaatttaagttgatttttttagcacatttcacg 1280
1281 taattaaataaaatattccagtatatttatatttatttccttacgttatctaaattttttgaagattaaaaacttttgatat_{R4} 1360

6/24

FIGURE 2-2

1361 aggcaagatcatgacacgctcgaagttaagtgaatgagactcctaacaaggtaataaacaagcagttcataaaaccggaatga 1440
 1441 ccttgatctttactaagcttgagatcattgaacatataaattaaatcgtttaatgaaagataagaactttaataataaaaat 1520
 1521 cattcaaaacgagaaactgataaacaacaaacgcaacgcaacaaataataatagacggtggaaggatgatgcagagcc 1600
 1601 atccaccctttttccagtttccctactgcttacttctctatgcataatcacaaagcgccttgaaacttgtagtcattg 1680
 1681 cagagcccttactcgccaggtcaccgaccacggtgttactctatcacttctcctccttcccttaaaagaaccaccacgc 1760
 1761 cacctccctctcacaacactcataaaaaaacccacctcttgcatcttcccaagttcaaaattagttcacagtaagcaag 1840
 1841 aactcaacaaca ATG GCG GAT CGT ACA ACA CAG CCA CAC CAA GTC CAG GTC CAC ACC CAG CAC 1903
 1 M A D R T T Q P H Q V Q V Q V H T Q H 17
 1904 CAC TAT CCC ACC GGC GGT TTC GGC CGT TAT GAA GGT GGA CTC AAA GGC GGT CCA CAT 1963
 18 H Y P T G G A F G R Y E G G L K G G P H 37
 1964 CAC CAG CAA GGA TCA GGC AGC GGC CCA TCA GCT TCC AAG GTG TTA GCA GTC ATG ACC GCG 2023
 38 H Q Q G S G S G P S A S K V L A V M T A 57
 2024 CTC CCC ATC GGC GGC ACC CTC CTT GCC TTG GCC GGG ATA ACC TTG GCT GGG ACG ATG ATC 2083
 58 L P I G G T L L A L A G I T L A G T M I 77
 2084 GGG CTG GCG ATC ACC ACC CCG ATT TTT GTC ATC TGC AGC CCT GTT CTA GTC CCG GCC GCT 2143
 78 G L A I T T P I F V I C S P V L V P A A 97
 2144 CTG CTC ATC GGG TTT GCC GTG AGC GCG TTT CTG GCC TCG GGG ATG GCC GGT CTA ACA GGG 2203
 98 L I G F A V S A F L A S G M A G L T G 117
 2204 CTG ACC TCG CTG TCG TGG TTT GCG AGG TAT CTG CAG CAG GCT GGG CAG GGA GTT GGA GTG 2263
 118 L T S L S W F A R Y L Q Q A G Q G V G V 137
 2264 GGG GTG CCG GAT AGT TTC GAG CAG GCG AAG AGG CGC ATG CAG GAT GCT GCT GGG TAT ATG 2323
 138 G V P D S F E Q A K R R M Q D A A G Y M 157

FIGURE 2-3

2324 GGG CAG AAG ACC AAG GAA GTT GGG CAG GAG ATC CAG AGG AAG TCT CAG GAT GTG AAA GCA 2383
158 G Q K T K E V G Q E I Q R K S Q D V K A 177

2384 TCA GAC AAA TAA ggtgataataaggggttttgggttcgtgtgtaaaactggtaaaaatctcgggtttttactg 2459
178 S D K * 181

2460 tacttttgcatgtagtgaatgagttcttgttctcttttgcctttttaatacataaagtaagaagcagcatattcatgt 2539

2540 tctggtgaatatgtcaagaattcgcaacaaatttagctaaccaggttcaatcttacoggtagacgacttcccagtaa 2619

2620 gaaacattccagggtccatcccggtataagagtctggacttctgaaaccttttagaccttggaatttgaaaaaagatgaaac 2699

2700 ctttagaataaatcacacgatggcagattgtacaaaactggagtcgagatcatgtaaattagccccataactaagaaccg 2779

2780 gcgatgacaacaattactaggaatatggttgtggcggtgtagcgggtgatttgaagaatcggggatcc 2859

2860 agaattgtgagaaccgaatcatcgacgaacattaccgccggaggagccattccaagcaactttggaactcctatatggct 2939

2940 gtccagcaggccacctgctcaagaaagaagccatgtcagaaaatccttaagaaaatcctaactggatgctgatatgaa 3019

3020 tccgccaggtgtgcggagttctttacaggcaggatctataaagaagaaacatgttttgtattggcattgttgatgttcca 3099

3100 agcacgcagcgatctatctccggtatcttaacaaacaaatacggattctctgaagaaacaaagcgcagaaaaacttctgcaac 3179

3180 gaaaccactcgtatatattgttctgagttggagaaagatgaccatactactctatttgggttgaaacttggaattggaaccga 3259

3260 aattttgagttgaaaaagcagtgatcgatatataaatttcagattcagattagatatcctatgagagaaggtagagttac 3339

3340 ctgatactactactgcccatcaggggtaaaaagtgcctcgatggttgttggagatggttccaggctaaatccacaa 3419

3420 cgctgaacaaattaaaaagatgaatggatcaatcttcaacccttacttctgcatttatgaggattggctcaaggctctcta 3499

3500 ga 3501

8/24

FIGURE 3-1

1	<u>tccactatgttaggtcatatcccatcttttaatttttgggcaccattcaattcccatcttgccttttagggatgtgaaataga</u>	5' primer (1)	AT rich	80
81	<u>acggccaaggtaagagaaataaaaaataatcccaaatataaagcaagagagggccaaagtaagataatccaaatgtacacttgtca</u>	AT rich		160
161	<u>tcgcccgaatattagtaaaatacgcggcatattgtatttccacacattattaaaaacogtataatgtattggctgcatttgc</u>			240
241	<u>atgaataatactacgtgtaagccccaaaagaacccacgtgttagcccatgcaaaagttaaacactcacgacccccattcctcagt</u>	RY	G box seed-specific	320
321	<u>ctccactatataaaacccaccatcccccaatctttaccaaaacccaccacagactcacaaactcgactctcacacaccttaaaagaa</u>		3' primer (1)	400
401	<u>ccaatcaccaccacaaaaATGGCAAAGCTGATGAGCCTAGCAGCCGTAGCAACGCAGTTCCTCTTCCTGATCGTGTGGAC</u>	TATA		480
1	<u>M A K L M S L A A V A T Q F L F L I V V D</u>			21
481	<u>GCATCCGTCCGAACCACAGTGATTATCGACGAGGAGACCAACCAAGGCCGGTGGAGGCAAGGTGGCAGGGACAGCAGC</u>			560
22	<u>A S V R T T V I I D E E T N Q G R G G K V A G T A A</u>			48
561	<u>AGTCTGCGAGCAGAGATCCAGCAGCGAGACTTCCTGAGGAGCTGCCAGCAGTTCATGTGGGAGAAAGTCCAGAGGGCGG</u>			640
49	<u>V C E Q Q I Q Q R D F L R S C Q Q F M W E K V Q R G G</u>			75
641	<u>GCCACAGCCACTATTACAACCAGGGCCGTGGAGGAGGCGAACAGAGCCAGTACTTCGAACAGCTGTTTGTGACGACCTTA</u>			720
76	<u>H S H Y Y N Q G R G G G E Q S Q Y F E Q L F V T T L</u>			101
721	<u>AGCAATTGCGCACCGGGTGCACCATGCCAGGGGACTTGAAGCGTGCCATCGGCCCAATGAGCAGGAAATCCAGCAGCA</u>			800
102	<u>S N C A P R C T M P G D L K R A I G Q Q M R Q E I Q Q Q</u>			128
801	<u>GGGACAGCAGCAGGACAGCAGGAGGAAAGTTACAGAGGTGGATCCAGCAAGCTAAACAAATCGCTAAGGACCTCCCCGGAC</u>			880
129	<u>G Q Q Q Q G Q Q Q E V Q R W I Q Q A K Q I A K D L P G Q</u>			155

10/24

FIGURE 4-1

10 20 30 40 50 60 70 80 90 100
 ctcaagcatacggacaaagggtaaataacatagtcaccagaaacataataaacaagtgagaagcaagataaaaaattagctatggacattcaggttc
 110 120 130 140 150 160 170 180 190 200
 atattggaaacatacattatcctagtgtgaccatccttccctcctgctctagttgagaggccttggaactaacgagaggtcagttgggtagcagatcc
 210 220 230 240 250 260 270 280 290 300
 ttatcctggactagcctttctgtgttcagagtcttcgtgccgcgtctacatctatctccattaggtctgaagatgactcttcacacccaacgacgttt
 310 320 330 340 350 360 370 380 390 400
 aaggctctatcctactcctagcttgcaatacctggcttgcaatacctggagcatcgtgcacgatgattggatactgtggaggaggtgtttgctgatt
 410 420 430 440 450 460 470 480 490 500
 tagagctcccggtgggtgatttgacttcgatttcagtttaggcttggtgaaatttttcagggtccattgtgaagccttagagcttgagcttccttcca
 510 520 530 540 550 560 570 580 590 600
 tgttaatgccttgatcgaaattctcctagagaaaggaagtcgatctctgagtatgaaatcgaaagtgacacatttttttcaacgtgtccaatcca
 610 620 630 640 650 660 670 680 690 700
 caaacaagaagaagacaggtaattcttccatactatactgacaagtaatagtcttaccgtcatgcataataaacgtctcgttccttcaagagggttttc
 710 720 730 740 750 760 770 780 790 800
 cgacatccataacgacccgaagcctcatgaaagcattagggaagaacttttggttcttctgtcatggccctttatagggtgtcagccgagctcgccaattc
 810 820 830 840 850 860 870 880 890 900
 ccgtccgactggctccgcaaaatattcgaaacggcaagttatggacttgcaaccataactccacggttattgagcaggacctattgtgaagactcatctcat
 910 920 930 940 950 960 970 980 990 1000
 ggagcttcagaatgtggtgtgtcagcaaaaccaatgaccgaaatccatcacatgacggacgtccagtggtgagcgaacgaagaaagcgcctatcttt
 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
 cagagtcgtgagctccacacccgattccggcaactacgtgttggcaggcttcgcgtattagagatatgttgaggcaagaccatctgtgccactcgta
 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 caattacgagaggtgtttttttgtgattttcttaagtttctcgttgatggtgagctcatattctacatcgatggtgtctctcaacgtgttctcgtgcat

12/24

FIGURE 4-3

2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
ACACCCCTCAGCTCATCTACATCGTTCAAGgtataaaatcagttcatacaaatgataaccaccacttcgaatgtatttatcaaatatcaatgatcga
N T P Q L I Y I V Q

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500
tgcacctgtatgtgtgtatattcagGTAGGGGAGTTACAGGAATCATGTTCCCAKATGTCCAGAGACATTCGAGGAATCCCAGCAAGGACAAC
G R G V T G I M F P X C P E T F E S Q Q G Q

2510 2520 2530 2540 2550 2560 2570 2580 2590 2600
AGGGCCAAACAGGGTAGTTCCTCCCAAGACAGCACAGAGATCCGCGCTTCGTTGAAGGTGACGTCAATGCCGTCCCTGCCGGTGTAGCCCACTGGTCCTA
Q G Q Q G S S Q D Q H Q K I R R F R E G D V I A V P A G V A H W S Y

2610 2620 2630 2640 2650 2660 2670 2680 2690 2700
CAACGATGGCAACGAACCAAGTCATGGCCATTGTTGTCCATGACACTCCAGCCACCTCAACCAACTGGACAACAACCCAGGgtatataagcattgcccgt
N D G N E P V M A I V V H D T S S H L N Q L D N N P R

2710 2720 2730 2740 2750 2760 2770 2780 2790 2800
agttgctaataaattgcacacaattggaaactctatttcagtatctataaacttttctcttttttgccagAACTTCTACTTGGCAGGAACCCGAGAGAC
N F Y L A G N P R D

2810 2820 2830 2840 2850 2860 2870 2880 2890 2900
- GAGTTCGAACAATCGCAGCAAGAGGCGAGGCTGAGCCGTGGGGAGAGTGAAGGTGACGAGGACGCGAGGGAACCTTCTCAACCTGCAACAACCTCTTCTT
E F E Q S Q Q G G R L S R G E S E G R G R R E P L Q P A T T S S

2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
GCGGAATCGACTCCAAGCTCATCGCGGAGGCGTTCAATGTGACGAGAACGTGGCAAGGAGGCTACAGAGCGGACGACACACAGAGGCCAGATCGTCCG
C G I D S K L I A E A F N V D E N V A R R L Q S E N D N R G Q I V R

3010 3020 3030 3040 3050 3060 3070 3080 3090 3100
AGTCGAAGGCGAGCTCGACATCGTCAGACCTCCGACCAAGTATCCAGGAGGAGTCAAGGAGGAGGTCGTGGTGGTGGCCGCTACTACTCCAATGGA
V E G E L D I V R P P T S I Q E E S Q E Q G G R G G R Y Y S N G

3110 3120 3130 3140 3150 3160 3170 3180 3190 3200
GTGGAGGAGACCTTCTGTCCATGAGACTAATTGAGACATCGGCGATCCTTCTCGGGCAGACATTTTCACTCCAGAAGCCGGCCGCTTAGATCCCTCA
V E E T F C S M R L I E N I G D P S R A D I F T P E A G R V R S L

13/24

FIGURE 4-4

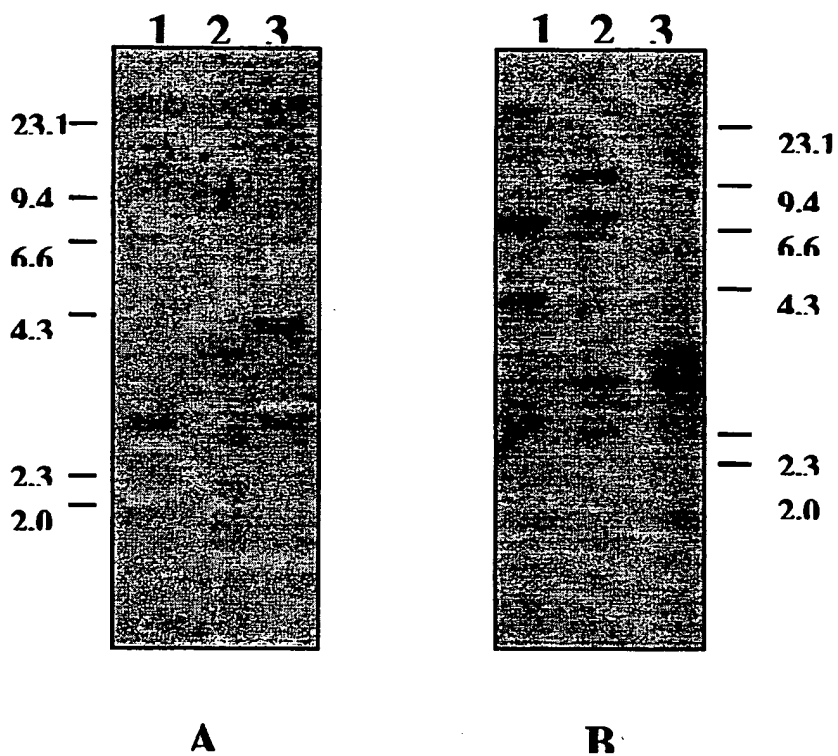
3210 3220 3230 3240 3250 3260 3270 3280 3290 3300
 ACAGCCAACTCCCGTCTGCAATGATCCAGCTTAGCGCGAGAGGCGTTCTCTACAAATgatatagatctcactcagcaccactctctaaattga
 N S H N L P V L Q W I Q L S A E R G V L Y N
 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400
 atccctaatttaattcaccgatattctgaccgaccggtttgaattttagAAGCGATCAGGTCGCCGACTGGAACATCAACGCACACAGCATAGT
 E A I R L P H W N I N A H S I V
 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500
 GTACGGATCAGAGGACAAAGCAGAGTCCAGATCGTGAACGAGGAAGGAATTCGGTGTTCGATGGAGTCTGCAGGAAGGACAGGTGGTGACGGTGCCG
 Y A I R G Q A R V Q I V N E E G N S V F D G V L Q E G Q V V T V P
 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
 CAGAACTTCGCGGTGTAAGAGATCCAGAGGAGAGGTTTGAAGTGGTGGCGTTCAAGACCAACGACACGCGATGGTGAACCTCGCTAGCCGGGAGGA
 Q N F A V V K R S Q S E R F E W V A F K T N D N A M V N S L A G R
 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
 CATCGGACTAAGGCGATCCCGCGGATGACTGGCTAACGCTGGAGGTTGCGCGAGGAGCGAGGAGGTGAAGTTCAACAGGCAGGAGACTCA
 T S A V R A I P A D V L A N A W R V S P E E A R R V K F N R Q E T H
 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800
 CTTGGCTAGCACCCAGGGCCAGTCCAGGTCGCCCGGAGGTGAATGTCGCAAGGAGGTGATCAACTTGTCTATGTAAaatgtgacggtgaaataataa
 L A S T R G Q S R S P G R L N V V K E V I N L L M *
 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900
 cggtaaaatatatgtaataataataataaagccacaaagtgagaatgaggggaaagtgtgtaatgagccagtagccggtggtgctaatttg
 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000
 tatcgtattgtcaataaatcatgaatttgggttttttatgtgttttttaaatcatgaatttttaaaattttataaaaaataatctccaatcggaagaacaac
 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100
 attccatatccatggatgttttctttaccctaaatcttagtttttgagaggatcaccgaacagttctgtgcaactatccctcaaaagctttaaaatga
 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
 acaacaagggaacagagcaacgttccaaagatcccaaacgaacatatattatctataactataattattataactactgcccgggaatcaccaatccct

14/24

FIGURE 4-5

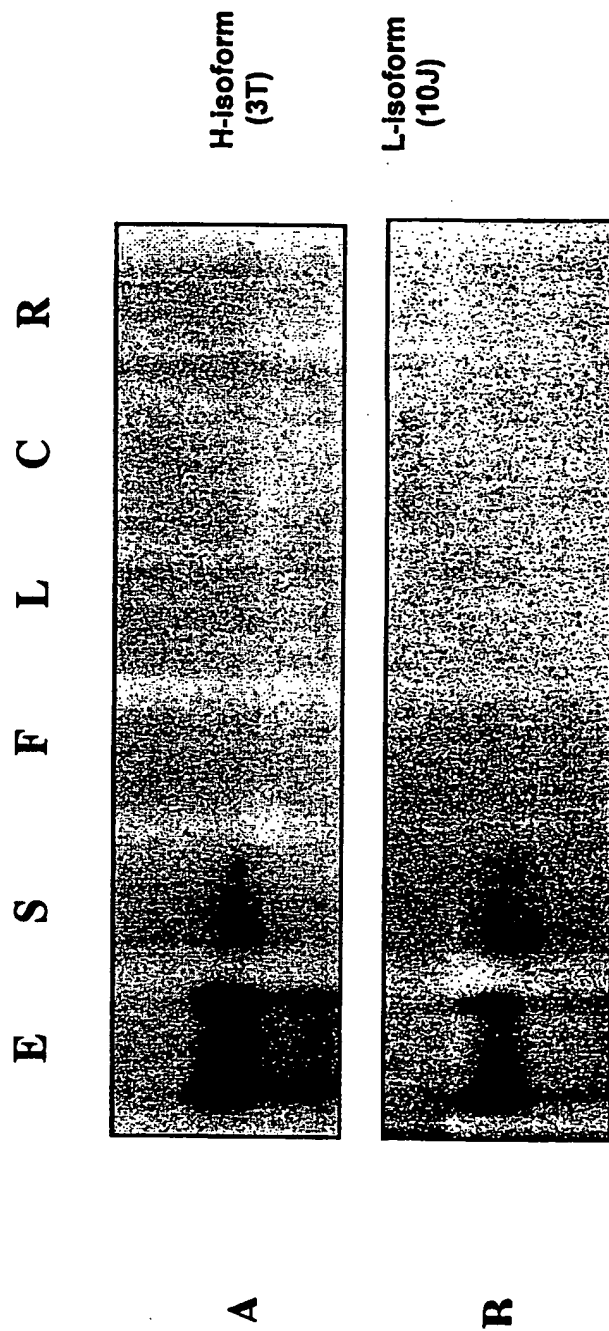
4210 4220 4230 4240 4250 4260 4270 4280 4290 4300
gaatgattcctattaactacaagccttggtggcgcgagagaagtgatcggcgcggcgagagcagcggaactcggagacgagccttgatgagcagagtc
4310 4320 4330 4340 4350 4360 4370 4380 4390 4400
tttacctgccaggcggtgaagggaagagcgcccttctggagtaggagttcagaaagcggttccttggcgagtaagcggaacgttaagggtggtgtgc
4410 4420 4430 4440 4450 4460 4470 4480 4490 4500
gacgtcctcggttcnggagcgcnattcatgaagggttaagtcanaatctgtagctctcgagtgctcagggagccnaaagacgttgggaaaccgtcgncgt
4510 4520 4530 4540 4550 4560 4570 4580 4590 4600
ttggggcatcagtcngcggggcaacgcttcctcctgctgctccanaancnangtanatttaaaaganatgggaaattaantaatggnnaatnannaggagg
4610 4620 4630 4640 4650 4660 4670 4680 4690 4700
attgnaacgggtcngancngnanganaganagtttttannggtttaataactgggggagtgngnagccngcctgttccngttagangaaaccaagnnccgg
4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
gaggnttncannngnnaggagagaaaagganncaatttnannangengaggagacatgaancggtacngagcgtgnggttcannnnnancggcgnnngnagtcc
4810 4820 4830 4840 4850 4860 4870 4880 4890 4900
cnngggaccnggntgggtgnanaagggaanggaacattnggtngnanganaaanaccttttacnattgccttgcaggnngtntngcncntncgggt
4910 4920 4930 4940 4950 4960 4970 4980 4990
nacatnccgctgcatgggcttggggngccnanaaggnagccncanggggnanncngccncttgtncangncgctnaagttcnattgtanattggnctgtt

15/24
FIGURE 5



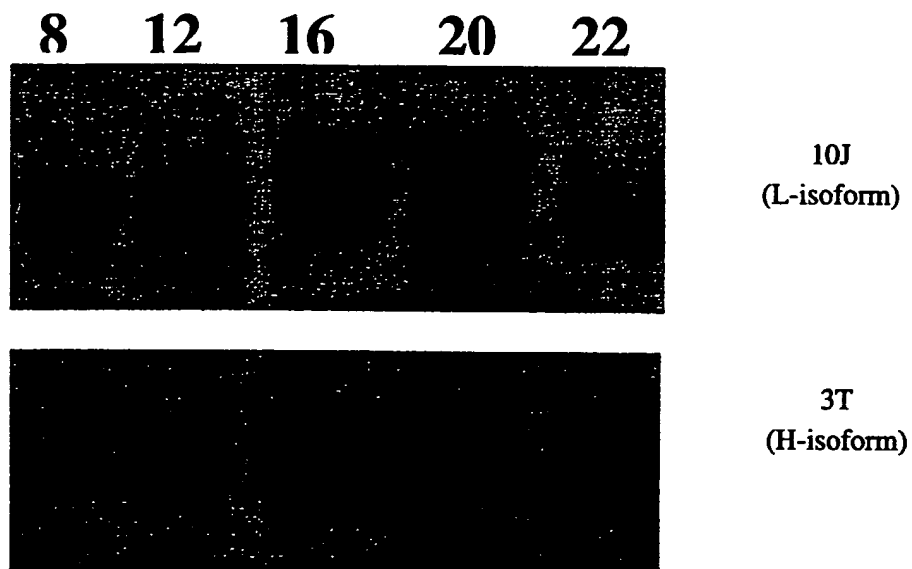
16/24

FIGURE 6

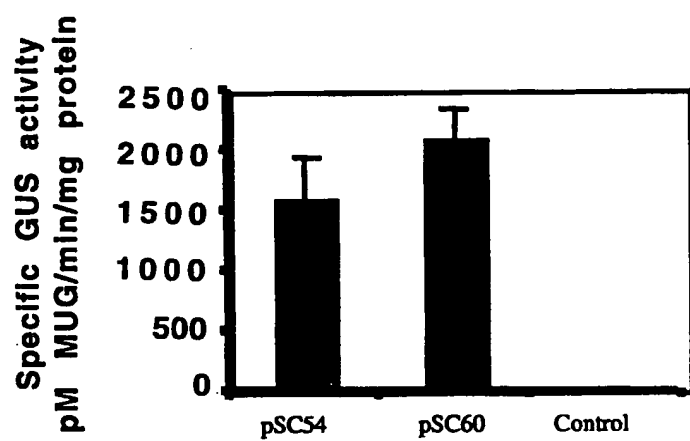


17/24

FIGURE 7



18/24
FIGURE 8



19/24
Figure 9.1



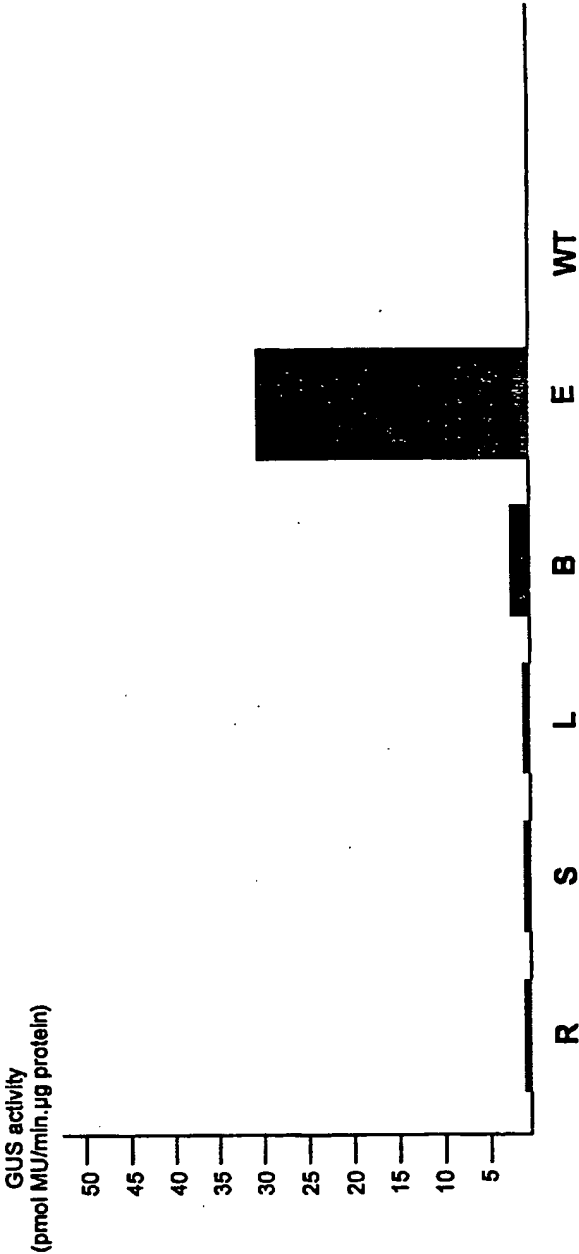
20/24

Figure 9.2



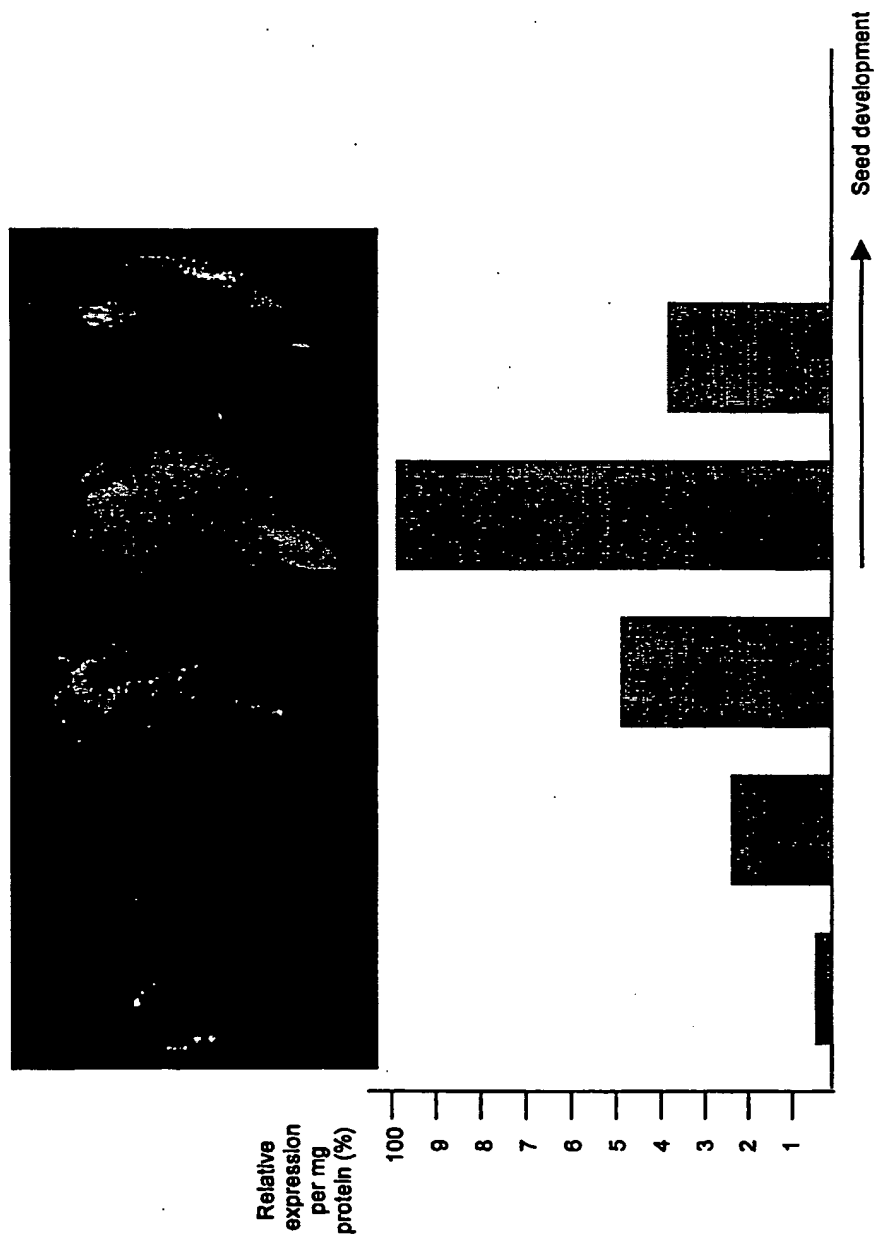
21/24

FIGURE 10



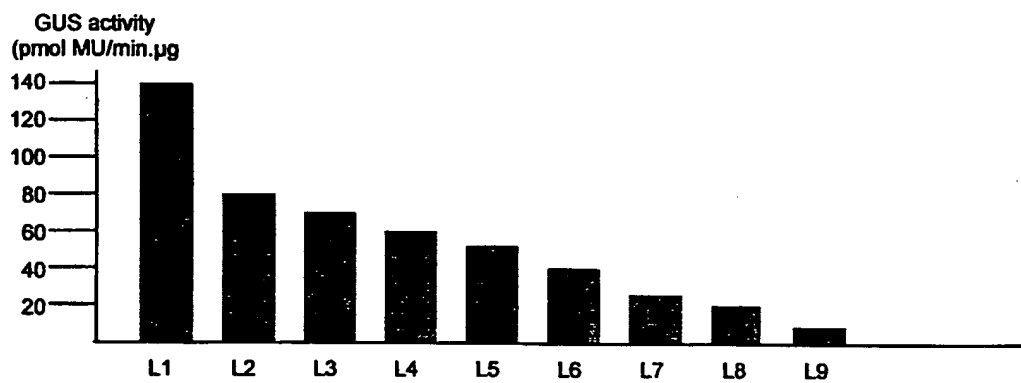
22/24

FIGURE 11



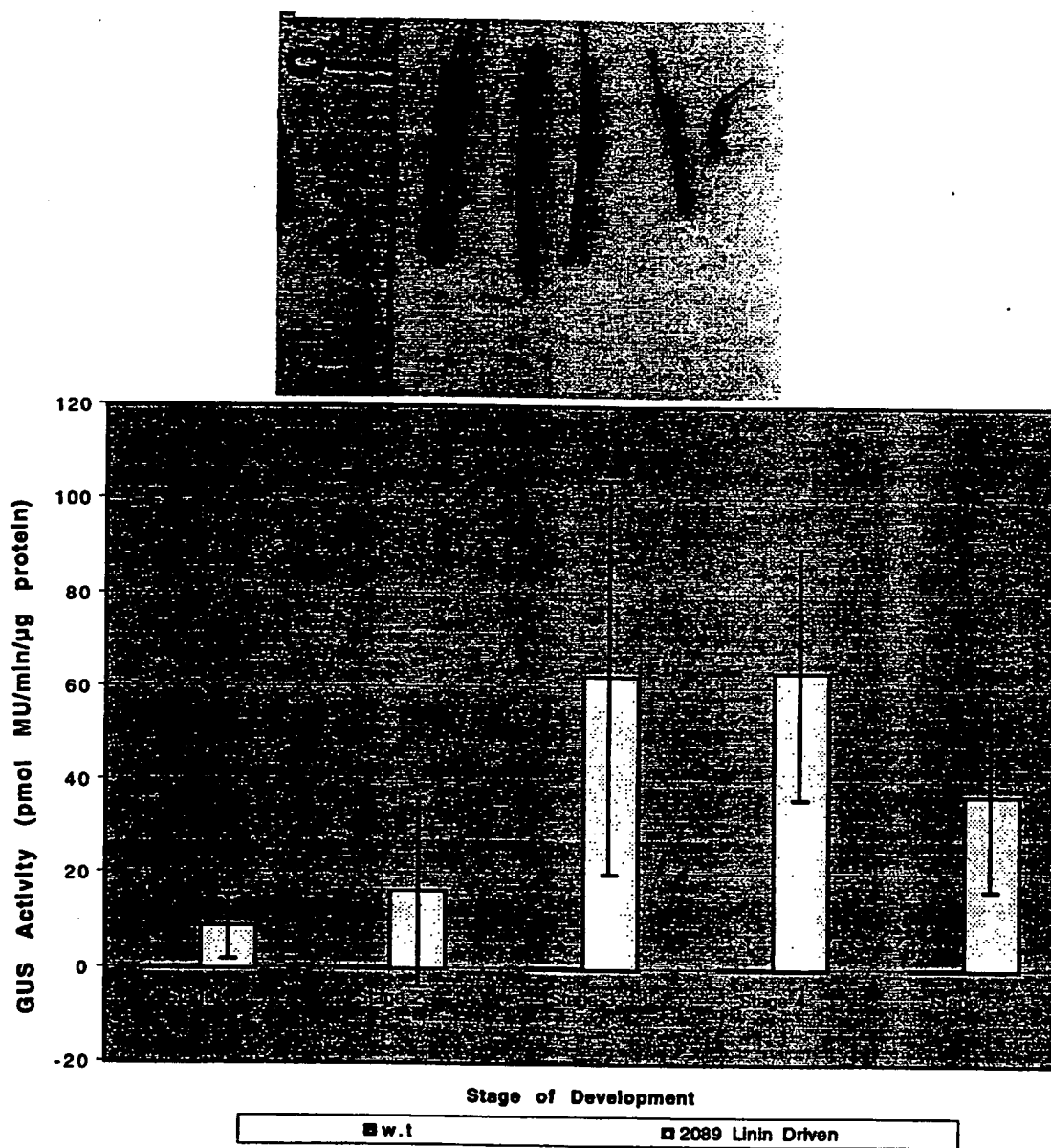
23/24

FIGURE 12



24/24

FIGURE 13



1/11

SEQUENCE LISTING

<110> Chaudhary, Sarita
van Rooijen, Gijs
Moloney, Maurice
Singh, Surinder
SemBioSys Genetics Inc.
Commonwealth Scientific and Industrial Research Organisation

<120> Flax Seed Specific Promoters

<130> 9369-147

<140>

<141>

<150> 60/151044

<151> 1999-08-27

<150> 60/161,722

<151> 1999-10-27

<160> 25

<170> PatentIn Ver. 2.0

<210> 1

<211> 4305

<212> DNA

<213> Linum usitatissimum

<400> 1

ttcaaaaccc	gattccccgag	gcggccctat	tgaagatatg	ggggaagttc	gacgagatcg	60
atgtcggggtc	gagtgctatg	gtgatgggtgc	cgtttggggg	gaggatgagc	gagatagcca	120
agactagcat	tccgttccca	cacagagttg	ggaatttgta	ccaaatccaa	cacttgctcg	180
attggagcga	cgatagggac	gcggaaaaac	acatccgttg	gatcagggag	ttgtacgatg	240
atctcgagcc	ttatgtgtcg	aagaatccga	ggtatgctta	cgtgaactac	agggatctcg	300
acatcgggat	gaatggagga	ggtgaagggg	atgagaaggg	tacttatggg	gaggctaagg	360
tgtgggggga	gaagtacttt	gggtcgaact	ttgatcgggt	ggttcgggtg	aagacgattg	420
ttgatcccaa	taatgtgttt	cgaaacgagc	agagcattcc	ctcaattcca	actcggttat	480
aaggatcaat	gtcaatgag	aattttcctt	tccaatgtga	ttacaagttc	tattgggtca	540
gctttctcaa	ctgctcctat	tcatttagat	taattcataa	caactattaa	tttaccagcc	600
ttttatccgg	cccgttggcc	gatttatatt	cttaagtttt	agatgaaatg	aaaccgattt	660
agtttttatt	gagatgagat	taactctaat	ttgcttgaaa	tttactcacg	gttgatgtga	720
tatttggaa	taactaaaat	gataaataat	ggataaaaaa	aaaaatattt	aaaataaata	780
acataaacat	aagaacaata	aaataaataa	atttaatttt	aatttatatt	cttgttttct	840
ttctgtatca	tacatctctt	ctcttacttc	ttaaaggctt	ttcaattatc	acttaattaa	900
atacaataga	taaactcgta	attctataac	attaacctat	acacttgcac	ggtgaacaat	960
caatatgata	atataataat	aatataataa	ttcaattatt	aatctacaat	tttttaatta	1020
taaagtttat	gcggtcagtt	tctgcaagct	ccgagctcct	tgtcatcggt	agtttctgcg	1080
gtctcaaggt	ataacgactc	ggagcgacga	gccctttgct	tccaatggac	gggttgcat	1140
tctgccgctg	ttgagctcga	ttggcgtgtc	atgctggagt	cagagttcct	acaaaaaac	1200
cctaaactag	agggtgatta	gggtgaaatt	agggtgttgg	cctgggttcc	attgtccaaa	1260
gttttagtca	acttaaaaac	agacttaaat	tttatgcttc	aaaatagttt	atctgttatt	1320
atattagcgt	gtaattagtc	ttgacaatgg	ggccggacgg	gtacggattc	gggacccga	1380
tccccgccca	tagtgtaatg	gctcaactgc	caagtcagca	ttggaccgaa	attattggac	1440
acgaagtact	aatgtgaaaa	actttacatt	tgttattttc	tactttaata	ctatgctatt	1500
ttcaaaattt	gaactttt	actatgtttt	tatatagttt	agtatatctt	aatttttatg	1560
caaattcatc	taattgtatt	aaactatttt	cgatccgtag	ctaattattt	cgaaggcaag	1620
tcaaagtgtt	attgtggact	atgtgagcta	atattgaacc	tttatctctc	ccaaccactc	1680
aagttaattg	aaccaaactc	gatcgggttg	gtttcgagct	atttcgagcc	attgtgttta	1740
tatgcacgtg	agatatcaag	attgacccga	acactttatt	atgataatgt	agaaaaagaa	1800
aacatattct	aagactacat	gcatgcaaag	tgcaaccctt	gcatggaaag	ctgctcaaca	1860
cgtggcatag	actcccgcca	cgtgtccatt	ccacctcatc	acctcaccac	caccgttcac	1920

2/11

```

ctcttattat atcacaacaa tcaatcaatc ctactcctcc atactcgaac aaatccgacc 1980
aacttatacc aatattccca aacttgatta atttctcagc aatatggatc agacgcacca 2040
gacatacgcc ggaaccacgc agaaccgcgag ctatggcggc gggggcacaa tgtaccagca 2100
gcagcagccg aggtcttacc aggcgggtgaa ggcggccact gcagccaccg cgggtggatc 2160
cctcatcggt ctgtccgggtc tcatccttac ggccaccgtc atttcaactca tcatagccac 2220
ccctctcctt gtcattctca gccctgttct tgtcccggtc ctcatcaccg tcgggtctctt 2280
gatcaccggg tttcttgctt ccggtgggtt cggagtcgcc gccgtcaccg tcttgtcctg 2340
gatctatagg tatgtataag ctttggactt tagtattgtt ataaaatata taagctgatt 2400
tatgaacatg gatctcccaa caagagttat ttaaattgcat tctcggtctg actcgatcgg 2460
ttgggttttg agctactcgg tcacaatggt cgggtcgggt ctggatctgt tatactaata 2520
tttggaagcc tgaagtttca ttgttctgcc ccaacttccc actacctttt gaggggtgta 2580
agaagccata caaactaatt atgaatccct cccaacaact cagaactcga gtcagtgggt 2640
tgtgaacatg ctctataaac atttcgaaaa tctttgttca atgaacgtag aaatgaccat 2700
gcttgatgat tgtgggtctt ataaggtacg tgaccggcgg gcaccggcg ggaggggatt 2760
cgctggacca ggctaggtcg aagctggccg gaaaggccag ggaggtgaag gacagggcgt 2820
cggagtccgc acagcagcat gtcacaggtg gtcaacagac ctcttaaaga gagtccctcta 2880
gttaaatattg tcttcgttct tgtttcgtgg cggcttgtaa actctctttt aagtgtgctg 2940
tttctctttt gtctcgtgtg ttgtaagtga aagtgtaatc gaagttccaa gttggagatg 3000
tttgtaacga tgatgttttc taataatcag agatattaaa agggttgcta atttagtatt 3060
gcgtctgatc tcggaccaaa ctcgcaagta aaattgcaga ggatgagttg tacagaacaa 3120
gcgtgcattg ttctggaagt tcatctcctt ggagccgacc ttgttgcttg cagtttcgcc 3180
aagtccacta gacaatgtta cgagttaagc ctctgtcaaa cagatcgctc tagcgtecca 3240
gaaaacacca gatttttcga aaaccatcgg ggatcaattt tcgattcaat tccgatcttg 3300
gaagtacttg aacagaagca tgatgctaaa agataataga aaatcgaagc ctagaaaagt 3360
tgtacagaaa gcaacaagtc aaaaatatag atcaacttca aaggttcaaa ttacatctta 3420
cagaccccaa aaaatgacag ttaacagaag tcgactaaac agaaaccagc cagcttcacc 3480
tggaatgaag gagctttgat caatccatcc tagcttcatt cccctttgaa attgcagaca 3540
gagctctcat cctgctaaag ctggtggctt attcttaacc ctgcaatcaa taagcatgaa 3600
ctaacattgg acaccttcat cggcggattg ctcgaaaaatc agtgagcgag ggatttacct 3660
gtgtgtgtag taacctctct ccttgatcat aaaatctgga aattccggca tcaactactg 3720
ccacctttct gcttaagggtg attttatcac caaggctgag cgtgattcct tgcgtcttg 3780
tccgaatcct gatgtatcca ctgagcttcc catctccttc cttctccagg cttatgttca 3840
ccaatgcgtc ctgcgccgaac acactcttgg cgtacaagtt cgcagccagg aatccacact 3900
ctccatcaag tgcagactg caaaccccaa ataagaacac aaactccaaa gtcaacgatc 3960
aattctccgc cttttatgaa gaaaaggaaa cttctgggta cttacgggtc cgtcagacac 4020
ttcatatttg tagacttgat gatattgtcc aggaattcct tctcgttctg aattgttgtg 4080
ttaacagcaa cctgacagac agaaagatat cgcaaattta agatactggg atgactaggc 4140
acagagaaat gaaatctaata tctagaagta aaaccttatt ttccatttca aattctgccc 4200
acatagtccg gaacgcagca tccgagcaag aagcaggaga gatgtaatcc atgatatcga 4260
tgtggatatt gttgaggacg acaactgaac gttccatcac attggg 4305

```

<210> 2
<211> 109
<212> PRT
<213> Linum usitatissimum

<400> 2
Met Asp Gln Thr His Gln Thr Tyr Ala Gly Thr Thr Gln Asn Pro Ser
1 5 10 15
Tyr Gly Gly Gly Gly Thr Met Tyr Gln Gln Gln Gln Pro Arg Ser Tyr
20 25 30
Gln Ala Val Lys Ala Ala Thr Ala Ala Thr Ala Gly Gly Ser Leu Ile
35 40 45
Val Leu Ser Gly Leu Ile Leu Thr Ala Thr Val Ile Ser Leu Ile Ile
50 55 60
Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Leu
65 70 75 80
Ile Thr Val Gly Leu Leu Ile Thr Gly Phe Leu Ala Ser Gly Gly Phe
85 90 95

3/11

Gly Val Ala Ala Val Thr Val Leu Ser Trp Ile Tyr Arg
 100 105

<210> 3
 <211> 46
 <212> PRT
 <213> Linum usitatissimum

<400> 3
 Tyr Val Thr Gly Gly His Pro Ala Gly Gly Asp Ser Leu Asp Gln Ala
 1 5 10 15

Arg Ser Lys Leu Ala Gly Lys Ala Arg Glu Val Lys Asp Arg Ala Ser
 20 25 30

Glu Phe Ala Gln Gln His Val Thr Gly Gly Gln Gln Thr Ser
 35 40 45

<210> 4
 <211> 3501
 <212> DNA
 <213> Linum usitatissimum

<400> 4
 tctagacatt tgacataaac cgaattcaaa gaacacaaca ttgactaaca ccaaaaagaa 60
 atagagtagt gaaatttggg agattaaaaa atagaaacaa actgattcctt agaaaagaaga 120
 gatgattagg tgctttcagt tcgggtctgtc agggaaatcga gatgttcact tatttacatt 180
 gtcgattcat ctoccaaattg tcctgggtcc ttactgtcc gacgcttttt tgaatcccag 240
 ttaattccca tcaagtcttc cttcagctgc gtagcactgc tagctccaac atggagcgtg 300
 gagtctactc gttcatgggg catcgcaaaag gtttgccttc atgttctgct accagccagc 360
 gccaccgcc tcttgggtgt gtggacaatt gcggtgaagc gcgcaagttg acatcccata 420
 gtctcgacac ttcaccatat ggatgtttaa aacgtatatc acgagtgcga tctacatgtc 480
 ccatcacacc acatataaag caatagtttg ggagcttttc atatttgaaa cgggcattga 540
 cgacttgccc tctcgataat ttaatctttt tttctcttca gctgattgtg tgcattccatt 600
 cgggctcaga agcacatcaa agggatctct ccatcgtagt attgggtcgt gtcgtatgat 660
 acgaagcagt cgatgaagtt tcctaagtgt cgagctacag gctccgcaaa gaaccgcga 720
 ggtagatcgt atgctagtag ccaaaaatca gtttgcgta gcggaatcaa cactagagac 780
 tcaccctaag gcatctcatg tgtgatgaac agtttatcat ttgtgagctt aggggtcatt 840
 gtcgatgacc caatgcacat tgagcttatg atagaatttg aataggaagc gttttccacc 900
 cagatcacga atagctaccc ctttttcggg cgccaaattt ccggcatcct atcttccacc 960
 acaacttaaa gatgcgatcg gtaaggaaact caccgaccac acacatcgaa taatcttcgg 1020
 tgaccgggtc ctgttgatca agtccctcaa tttcctcaac ctagtcttca atcgccgcta 1080
 gcgttatccc ccgcatatgg actttcatag cgcggagcgt agccggagac gacgagcaag 1140
 aaggatgagc ggccgagatg tgccggctaaa gaaacgagct tcctgccttg ctctatggag 1200
 gcagatttct gagttgatgg tgatggattt gtgatgtgga cacttttaac ttaagttgat 1260
 tttttagcac ttcatctcac taattaaata aataatttcc agtattttat atttatttcc 1320
 ttacgttatc taattttttg aaagattaaa actttgatat aggcaagatc atgacacgtc 1380
 gaagttaagt gaatgagact cctaacaagg taataacaaa gcagttcata aaccgaatga 1440
 ccttgatctt tactaagctt gagatcattg aacatataat taaatacggt aatgaaagat 1500
 aagaacttta atataaaaaa cattcaaaac gagaaactga taacaaaaac aaagcaaagc 1560
 gccaaacaaa taatagacgg tgggaaggatg atgcagagcc atccaccctt ttttcccagt 1620
 ttccttactg cttacttctc tatgcatatc acaagacgcc cttgaaactt gttagtcatt 1680
 cagagccctt actcgcagg tcaccgcacc acgtgttact ctatcacttc tcctcccttt 1740
 cctttaaaga accaccagc cacctccctc tcataaaaaa accacctctt 1800
 gcatttctcc caagttaaaa ttagttcaca gctaagcaag aactcaacaa caatggcgga 1860
 tcgtacaaca cagccacacc aagtccaggt ccacacccag caccactatc ccaccggcgg 1920
 ggctttcggc cggttatgaag gtggactcaa aggcgggtcca catcaccagc aaggatcagg 1980
 cagcgggcca tcagcttcca aggtgttagc agtcatgacc gcgctcccca tcggcgggac 2040
 cctccttgcc ttggccggga taaccttgcc tgggacgatg atcgggctgg cgatcaccac 2100
 cccgattttt gtcattctga gccctgttct agtcccggcc gctctgctca tcgggtttgc 2160
 cgtgagcgcg tttctggcct cggggatggc cgggctgaca gggctgacct cgctgtcgtg 2220
 gtttgcgagg tatctgcagc aggtcgggga gggagtggga gtgggggtgc cggaatgttt 2280
 cgagcaggcg aagaggcgca tgcaggatgc tgctgggtat atggggcaga agaccaagga 2340

4/11

```

agttgggcag gagatccaga ggaagtctca ggatgtgaaa gcatcagaca aataaggtga 2400
taataagggg ttttgggttc gtgtgtaaac tggtaaaatg gaaattcttg gttttactgt 2460
acttttgcat gtagtggaat gaatgagttc ttgttctctt ttgtctttta atcataaagt 2520
aagaagcagc atttcatgtt ctggttgaat attgtcaaga attcgcaaca aatttagcta 2580
aaccagttca atcttaccgg ttgacgact tcccagtaag aaacattcca ggtccatccc 2640
ggtataagag tctggacttc tgaacacctt agaccttga tttggaaaaa agatgaaacc 2700
tttagaataa attacaacga tggcagattg tacaaaaactg gagtcgagat catgtaaatt 2760
agcccataac taagaaccgg cgatgacaac aattactagg aatatgggtg ttgggctggt 2820
cggcgctag cggtgatgat ttggaagaat cggggatcca gaatgtgaga accgaatcat 2880
cgacgaacat taccggcgga ggagcccat tcaagcaact ttggaactcc tatatggctg 2940
ttccagcagg ccacctgctc aagaaagaaa gaagccatgt cagaaatcct tacgaaatct 3000
aactggatgc tgatatgaat ccgccaggtg tgcggagttc tttacaggca ggatctataa 3060
agaagaaaca tggtttgtat tggcattgtt cggtttccaa gcacgcagcg atctatctcc 3120
ggatcctaac aacaaaaata cggattctgt aagaaacaag cgcagaaaac ttctgcaacg 3180
aaaccactcg tatatttgt tctgagttgg agaaagatga ccatactact gtatttggtt 3240
gaacttggat tggaaaccgaa attttgagtt gaaaagcgag tgatcgtata taaatttcag 3300
attcagatta ggatattccta tgagagaagg tagagttacc tgatactaca tactgcccac 3360
caggggtaaa agttgcctcg atggttgtgt ttggagatgg ttccaggcta aatccacaac 3420
gctgaacaaa ttaaaagatg aatggatcaa tcttcaaccc ttacttctgc atttatgagg 3480
attggtctca ggctctctag a 3501

```

<210> 5

<211> 180

<212> PRT

<213> Linum usitatissimum

<400> 5

```

Met Ala Asp Arg Thr Gln Pro His Gln Val Gln Val His Thr Gln
  1             5             10             15

```

```

His His Tyr Pro Thr Gly Gly Ala Phe Gly Arg Tyr Glu Gly Gly Leu
          20             25             30

```

```

Lys Gly Gly Pro His His Gln Gln Gly Ser Gly Ser Gly Pro Ser Ala
      35             40             45

```

```

Ser Lys Val Leu Ala Val Met Thr Ala Leu Pro Ile Gly Gly Thr Leu
      50             55             60

```

```

Leu Ala Leu Ala Gly Ile Thr Leu Ala Gly Thr Met Ile Gly Leu Ala
      65             70             75             80

```

```

Ile Thr Thr Pro Ile Phe Val Ile Cys Ser Pro Val Leu Val Pro Ala
          85             90             95

```

```

Ala Leu Leu Ile Gly Phe Ala Val Ser Ala Phe Leu Ala Ser Gly Met
      100             105             110

```

```

Ala Gly Leu Thr Gly Leu Thr Ser Leu Ser Trp Phe Ala Arg Tyr Leu
      115             120             125

```

```

Gln Gln Ala Gly Gln Gly Val Gly Val Gly Val Pro Asp Ser Phe Glu
      130             135             140

```

```

Gln Ala Lys Arg Arg Met Gln Asp Ala Ala Gly Tyr Met Gly Gln Lys
      145             150             155             160

```

```

Thr Lys Glu Val Gly Gln Glu Ile Gln Arg Lys Ser Gln Asp Val Lys
          165             170             175

```

```

Ala Ser Asp Lys
          180

```

<210> 6

5/11

<211> 1676
 <212> DNA
 <213> Linum usitatissimum

<400> 6
 tccactatgt aggtcatatc catcatttta atttttgggc accattcaat tccatcttgc 60
 ctttagggat gtgaatatga acggccaagg taagagaata aaaataatcc aaattaaagc 120
 aagagaggcc aagtaagata atccaaatgt acacttgta tgcgccaaat tagtaaaata 180
 cgcgccatat tgtattccca cacattatta aaataccgta tatgtattgg ctgcatttgc 240
 atgaataata ctacgtgtaa gcccaaaaga acccacgtgt agcccatgca aagttaacac 300
 tcacgacccc attcctcagt ctccactata taaaccacac atccccaatc ttaccaaacc 360
 caccacagca ctcaacaact gactctcaca ccttaaagaa ccaatcacca ccaaaaaatg 420
 gcaaagctga tgagcctagc agccgtagca acgcagtcc tcttctgat cgtgggtggc 480
 gcatccgtcc gaaccacagt gattatcgac gaggagacca accaaggccg cgggtggaggc 540
 aaggtggcag ggacagcagc agtctgcgag cagcagatcc agcagcgaga cttcctgagg 600
 agctggccag agttcatgtg ggagaaagtc cagaggggag gccacagcca ctattacaac 660
 cagggcccggtg gaggaggcga acagagccag tacttcgaac agctgtttgt gacgacctta 720
 agcaattgctg caccgcgggtg caccatgcca ggggacttga agcgtgccat cggccaaatg 780
 aggcagggaaa tccagcagca gggacagcag cagggacagc agcaggaagt tcagaggtgg 840
 atccagcaag ctaaacaagt cgtaaggac ctccccggac agtgccgcac ccagcctagc 900
 caatgccagt tccagggcca gcagcaatct gcatggtttt gaagggttga tgcattatga 960
 gatcgtaaca agacactgct aggtgttaag gatggataat aataataata atgagatgaa 1020
 tgtgttttaa gttagtgtaa cagctgtaat aaagagagag agagagagag agagagagag 1080
 agagagagag agagagagag agaggctgat gaaatgttat gtatgtttct tggtttttaa 1140
 aataaatgaa agcacatgct cgtgtggttc tatcgaatta ttcggcggtt cctgtgggaa 1200
 aaagtccaga agggcgggccg cagctactac tacaaccaag gccgtggagg agggcaacag 1260
 agccagcact tcgatagctg ctgcgatgat cttaagcaat tgaggagcga gtgcacatgc 1320
 aggggactgg agcgtgcaat cggccagatg aggcaggaca tccagcagca gggacagcag 1380
 caggaagtgt agagggtggt ccatcaatct aaacaagtgc ctaggacct tccgggacag 1440
 tgccgcaccc agcctagccg atgccagctc caggggcagc agcagtctgc atggttttga 1500
 agtgggtgatc gatgagatcg tataaagaca ctgctagggtg ttaaggatgg gataataaga 1560
 tgtgttttaa gtcattaacc gtaataaaaa gagagagagg ctgatggaat gttatgtatg 1620
 tatgtttctt ggttttttaa attaaatgga aagcacatgc tcgtgtgggt tctatc 1676

<210> 7
 <211> 174
 <212> PRT
 <213> Linum usitatissimum

<400> 7
 Met Ala Lys Leu Met Ser Leu Ala Ala Val Ala Thr Gln Phe Leu Phe
 1 5 10 15
 Leu Ile Val Val Asp Ala Ser Val Arg Thr Thr Val Ile Ile Asp Glu
 20 25 30
 Glu Thr Asn Gln Gly Arg Gly Gly Gly Lys Val Ala Gly Thr Ala Ala
 35 40 45
 Val Cys Glu Gln Gln Ile Gln Gln Arg Asp Phe Leu Arg Ser Cys Gln
 50 55 60
 Gln Phe Met Trp Glu Lys Val Gln Arg Gly Gly His Ser His Tyr Tyr
 65 70 75 80
 Asn Gln Gly Arg Gly Gly Gly Glu Gln Ser Gln Tyr Phe Glu Gln Leu
 85 90 95
 Phe Val Thr Thr Leu Ser Asn Cys Ala Pro Arg Cys Thr Met Pro Gly
 100 105 110
 Asp Leu Lys Arg Ala Ile Gly Gln Met Arg Gln Glu Ile Gln Gln Gln
 115 120 125
 Gly Gln Gln Gln Gly Gln Gln Gln Glu Val Gln Arg Trp Ile Gln Gln

6/11

130

135

140

Ala Lys Gln Ile Ala Lys Asp Leu Pro Gly Gln Cys Arg Thr Gln Pro
 145 150 155 160

Ser Gln Cys Gln Phe Gln Gly Gln Gln Ser Ala Trp Phe
 165 170

<210> 8

<211> 4999

<212> DNA

<213> Linum usitatissimum

<400> 8

ctcaagcata	cggacaagg	taaataacat	agtcaccaga	acataataaa	caaaaagtgc	60
agaagcaaga	taaaaaaatt	agctatggac	attcagggttc	atattggaaa	catcattatc	120
ctagtcttgt	gaccatcctt	cctcctgctc	tagttgagag	gccttgggac	taacgagagg	180
tcagttggga	tagcagatcc	ttatcctgga	ctagcctttc	tggtgtttca	gagtcttcgt	240
gccgccgtct	acatctatct	ccattaggtc	tgaagatgac	tcttcacacc	aacgacgttt	300
aaggtctcta	tcctactcct	agcttgcaat	acctggcttg	caataacctgg	agcatcgtgc	360
acgatgattg	gatactgtgg	aggaggagtg	tttgcgtgatt	tagagctccc	ggttgggtga	420
tttgacttcg	atttcagttt	aggcttggtg	aaatttttca	ggttccattg	tgaagccttt	480
agagcttgag	cttcctttcca	tgtaaatgcc	ttgatcgaat	tctcctagag	aaaagggaag	540
tcgatctctg	agtaattgaaa	tcgaagtgc	catttttttt	caacgtgtcc	aatcaatcca	600
caaacaaagc	agaagacagg	taatctttca	tacttatact	gacaagtaat	agtcttaccg	660
tcatgcataa	taacgtctcg	ttccttcaag	aggggttttc	cgacatccat	aacgacccga	720
agcctcatga	aagcattagg	gaagaacttt	tggttcttct	tgatcatggc	tttataggtg	780
tcagccgagc	tcgccaattc	ccgtccgact	ggctccgcaa	aatattcgaa	cggcaagtta	840
tggagttgca	accataactc	cacgggtattg	agcaggacct	attgtgaaga	ctcatctcat	900
ggagcttcag	aatgtgggtg	tcagcaaacc	aatgaccgaa	atccatcaca	tgacggacgt	960
ccagtgggtg	agcgaaacga	aacaggaagc	gcctatcttt	cagagtcgtg	agctccacac	1020
cggattccgg	caactacgtg	ttgggcaggc	ttcgccgtat	tagagatatg	ttgaggcaag	1080
acccatctgt	ccaattacgag	caattacgag	agttgttttt	tttgtgattt	tcctaagttt	1140
ctcgttgatg	gtgagctcat	attctacatc	gtatgggtctc	tcaacgtcgt	ttcctgtcat	1200
ctgatatccc	gtcatttgca	tccacgtgcg	ccgcctcccg	tgccaagtcc	ctaggtgtca	1260
tgcacgccaa	attgggtggg	gtgcgggctg	ccctgtgctt	cttaccgatg	gggtggaggtt	1320
gagtttgggg	gtctccgcgg	cgatggtagt	gggttgacgg	tttgggtggg	ggttgacggca	1380
ttgatcaatt	tacttcttgc	ttcaaattct	ttggcagaaa	acaattcatt	agattagaac	1440
tggaaaccag	agtgatgaga	cggattaagt	cagattccaa	cagagttaca	tctcttaaga	1500
aataatgtaa	cccctttaga	ctttatatat	ttgcaattaa	aaaaataatt	taacttttag	1560
actttatata	tagttttaat	aactaagttt	aaccactcta	ttatttatat	cgaaactatt	1620
tgtatgtctc	ccctctaaat	aaacttggtg	ttgtgtttac	agaacctata	atcaaatatt	1680
caatactcaa	ctgaagtttg	tgcagttaat	tgaagggtat	aacggccaaa	atgcactagt	1740
attatcaacc	gaatagattc	acactagatg	gccatttcca	tcaatatcat	cgccgttctt	1800
cttctgtcca	catatccgct	ctgaaacttg	agagacacct	gcacttcatt	gtccttatta	1860
cgtgttacaa	aatgaaaccc	atgcatccat	gcaaactgaa	gaatggcgca	agaacccttc	1920
ccctccattt	cttatgtggc	gaccatccat	ttcaccatct	cccgtataaa	aacaccccca	1980
tcaactcacc	tagaacatca	tcactacttg	cttatccatc	caaaagatac	ccaccatggc	2040
tagatcatca	agccctttgc	ttctctcact	ctgcattttc	gccattctct	tccactcttc	2100
tctgggtagg	cagcaattcc	agcaggggaa	cgagtggccag	atcgacagga	tcgacgcac	2160
cgagccggac	aaaaccatcc	aggcagaagc	tggcaccatc	gaggtatggg	accagaaccg	2220
ccagcaattc	cagtgcgctg	gtgttgccgt	tgtaaggcgc	accattgagc	ccaaagggtc	2280
tctcttgcc	ttctacagca	acacccctca	gtcatctac	atcggttcaag	gtataaatta	2340
aatcagttca	tacaattgata	accaccactt	cgaatgtatt	tatcaaatat	caatgatcga	2400
tgcacctgta	tggtgtgtgt	atattcaggt	aggggagtta	caggaatcat	gttcccakga	2460
tgtccagaga	cattcgagga	atcccagcag	caaggacaac	agggccaaca	gggtagtcc	2520
caagaccagc	accagaagat	ccgccgcttc	cgtgaagggtg	acgtcattgc	cgctccctgcc	2580
gggtgtggcc	actggctcta	caacgatggc	aacgaaccag	tcatggccat	tggtgtccat	2640
gacacttcca	gccacctcaa	ccaactggac	aacaacccca	gggtatataa	gcattgcccgt	2700
agtttgcta	aaattgcaca	caattggaac	tctattttca	gtatctaata	actttttcct	2760
tttttggcag	aacttctact	tggcagggaaa	cccagagagac	gagttcgaac	aatcgacagca	2820
aggaggcagg	ctgagcagg	gggagagtga	aggtggacga	ggacgcaggg	aacctcttca	2880
acctgcaaca	acctcttctt	gcggaatcga	ctccaagctc	atcgcgagg	cgttcaatgt	2940
cgacgagaa	gtggcaagga	ggctacagag	cgagaacgac	aacagaggcc	agatcgctccg	3000

7/11

```

agtcgaaggc gagctcgaca tcgtcagacc tccgaccagt atccaggagg agtcacagga 3060
gcagggaggt cgtgggtggtg gccgctacta ctccaatgga gtggaggaga ccttctgctc 3120
catgagacta attgagaaca tcggcgatcc ttctcgggca gacattttca ctccagaagc 3180
cgcccgcggt agatccctca acagccacaa cctccccgtc ctgcaatgga tccagcttag 3240
cgccgagaga ggcggttctct acaatgtata gatctcactc acgcaccaac tctaaattga 3300
atccctaatt atttaattca ccgatatctg accgaccggt ttgaattttg taggaagcga 3360
tcaggctgcc gcactggaac atcaacgcac acagcatagt gtacgcgatc agaggacaag 3420
ccagagtcca gatcgtgaac gaggaaggga attcgggtgt cgatggagtg ctgcaggaag 3480
gacaggtggt gacggtgccg cagaacttcg cgttggttaa gagatcccag agcgagaggt 3540
ttgagtgggt ggcggttcaag accaacgaca acgcgatggt gaactcgcta gccgggagga 3600
catcggcagt aagggcgatc cccgcggatg tactggctaa cgctggagg gtgtcgccgg 3660
aggaggcgag gaggggtgaag ttcaacaggg aggagactca cttggctagc accaggggcc 3720
agtcagggtc gcccgggagg ttgaatgtcg tcaaggaggt gatcaacttg cttatgtaaa 3780
atgtgacggt gaaataataa cggtaaaata tatgtaataa taataataat aaagccacaa 3840
agtgagaatg aggggaaggg gaaatgtgta atgagccagt agccgggtgt gctaattttg 3900
tatcgtattg tcaataaatc atgaattttg tggtttttat gtgttttttt aaatcatgaa 3960
ttttaaatgt tataaaataa tctccaatcg gaagaacaac attccatata catggatgtt 4020
tctttaccca aatctagttc ttgagaggat gaagcatcac cgaacagttc tgcaactata 4080
cctcaaaagc tttaaaatga acaacaagga acagagcaac gttccaaaga tcccaaagca 4140
aacatattat ctataactat actatattat taattactac tgcccgggaat cacaatccct 4200
gaatgatttc tattaaacta aagccttgtt gccgcgagg aagtgatcgg cgcgcgagga 4260
agcagcgagc tcgggagacga ggccttggat gagcagagtc ttacctgcc agggcgtaga 4320
ggggaagagc ggccttcttg agtaggagtt cagcaagcgg cggttccttg gcggagtaag 4380
cggacgtaag ggtgntgtc gacgtcntcg tttcnggagg cgnattcatg aagggttaa 4440
gtcanatctg tagctctcga gtgctcaggg agccnaaaga cgttgggaaa ccgtcgnctg 4500
ttggggcatc agtcnccggg gcacgcttcc ctccgtgctc tccanaancn angtanattt 4560
aaaaganatg ggaaattaan taatggnaat nannaggagg attgnaacgg tcnganccgn 4620
angaanagtt tttannggtt taaatactgg gggagtngna gccngccnct ggttccngtg 4680
tagangaaac caagnnccgg gaggnntnca nnnngnaggg agaaaaagga nncatttnan 4740
nangcnaggg gacatgaanc ggtacngagc tgnnggttcan nnancggcgn nnggnagtcc 4800
cnngggaccn ggntggggtn anaagggaan ggaacattng gtngnangga naanaccntt 4860
ttacnattgc ctttgagagn nngtntnggc ncntnccggg nacatnccgc tgcattgggt 4920
ttggggngcc nanaggnagc cncangggna nncngccncc ttgtncangn cgctnaagtt 4980
cnattgtana tggncgttg

```

```

<210> 9
<211> 96
<212> PRT
<213> Linum usitatissimum

```

```

<400> 9
Met Ala Arg Ser Ser Ser Pro Leu Leu Leu Ser Leu Cys Ile Phe Ala
  1             5             10             15
Ile Leu Phe His Ser Ser Leu Gly Arg Gln Gln Phe Gln Gln Gly Asn
          20             25             30
Glu Cys Gln Ile Asp Arg Ile Asp Ala Ser Glu Pro Asp Lys Thr Ile
          35             40             45
Gln Ala Glu Ala Gly Glu Val Trp Asp Gln Asn Arg Gln Gln Phe Gln
          50             55             60
Cys Ala Gly Val Ala Val Val Arg Arg Thr Ile Glu Pro Lys Gly Leu
          65             70             75             80
Leu Leu Pro Phe Tyr Ser Asn Thr Pro Gln Leu Ile Tyr Ile Val Gln
          85             90             95

```

```

<210> 10
<211> 85
<212> PRT
<213> Linum usitatissimum

```

8/11

<400> 10

Gly Arg Gly Val Thr Gly Ile Met Phe Pro Xaa Cys Pro Glu Thr Phe
 1 5 10 15
 Glu Glu Ser Gln Gln Gln Gly Gln Gln Gly Ser Ser Gln
 20 25 30
 Asp Gln His Gln Lys Ile Arg Arg Phe Arg Glu Gly Asp Val Ile Ala
 35 40 45
 Val Pro Ala Gly Val Ala His Trp Ser Tyr Asn Asp Gly Asn Glu Pro
 50 55 60
 Val Met Ala Ile Val Val His Asp Thr Ser Ser His Leu Asn Gln Leu
 65 70 75 80
 Asp Asn Asn Pro Arg
 85

<210> 11

<211> 165

<212> PRT

<213> Linum usitatissimum

<400> 11

Asn Phe Tyr Leu Ala Gly Asn Pro Arg Asp Glu Phe Glu Gln Ser Gln
 1 5 10 15
 Gln Gly Gly Arg Leu Ser Arg Gly Glu Ser Glu Gly Gly Arg Gly Arg
 20 25 30
 Arg Glu Pro Leu Gln Pro Ala Thr Thr Ser Ser Cys Gly Ile Asp Ser
 35 40 45
 Lys Leu Ile Ala Glu Ala Phe Asn Val Asp Glu Asn Val Ala Arg Arg
 50 55 60
 Leu Gln Ser Glu Asn Asp Asn Arg Gly Gln Ile Val Arg Val Glu Gly
 65 70 75 80
 Glu Leu Asp Ile Val Arg Pro Pro Thr Ser Ile Gln Glu Glu Ser Gln
 85 90 95
 Glu Gln Gly Gly Arg Gly Gly Gly Arg Tyr Tyr Ser Asn Gly Val Glu
 100 105 110
 Glu Thr Phe Cys Ser Met Arg Leu Ile Glu Asn Ile Gly Asp Pro Ser
 115 120 125
 Arg Ala Asp Ile Phe Thr Pro Glu Ala Gly Arg Val Arg Ser Leu Asn
 130 135 140
 Ser His Asn Leu Pro Val Leu Gln Trp Ile Gln Leu Ser Ala Glu Arg
 145 150 155 160
 Gly Val Leu Tyr Asn
 165

<210> 12

<211> 141

<212> PRT

<213> Linum usitatissimum

9/11

<400> 12

Glu Ala Ile Arg Leu Pro His Trp Asn Ile Asn Ala His Ser Ile Val
 1 5 10 15
 Tyr Ala Ile Arg Gly Gln Ala Arg Val Gln Ile Val Asn Glu Glu Gly
 20 25 30
 Asn Ser Val Phe Asp Gly Val Leu Gln Glu Gly Gln Val Val Thr Val
 35 40 45
 Pro Gln Asn Phe Ala Val Val Lys Arg Ser Gln Ser Glu Arg Phe Glu
 50 55 60
 Trp Val Ala Phe Lys Thr Asn Asp Asn Ala Met Val Asn Ser Leu Ala
 65 70 75 80
 Gly Arg Thr Ser Ala Val Arg Ala Ile Pro Ala Asp Val Leu Ala Asn
 85 90 95
 Ala Trp Arg Val Ser Pro Glu Glu Ala Arg Arg Val Lys Phe Asn Arg
 100 105 110
 Gln Glu Thr His Leu Ala Ser Thr Arg Gly Gln Ser Arg Ser Pro Gly
 115 120 125
 Arg Leu Asn Val Val Lys Glu Val Ile Asn Leu Leu Met
 130 135 140

<210> 13

<211> 11

<212> PRT

<213> Linum usitatissimum

<400> 13

Gln Gln Gln Gly Gln Gln Gln Gly Gln Gln Gln
 1 5 10

<210> 14

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 14

tccactatgt aggtcata

18

<210> 15

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 15

ctttaagggtg tgagagtc

18

<210> 16

<211> 15

<212> DNA

<213> Artificial Sequence

10/11

<220>

<223> Description of Artificial Sequence: Primer

<400> 16

aggggtgatc gatta

15

<210> 17

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 17

gatagaaccc acacgagc

18

<210> 18

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 18

tatctagact caagcatacg gacaagggt

29

<210> 19

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: XbaI site

<400> 19

tctaga

6

<210> 20

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 20

ggttatcatt gtatgaactg a

21

<210> 21

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: NcoI site

<400> 21

ccatgg

6

<210> 22

<211> 32

<212> DNA

<213> Artificial Sequence

11/11

<220>

<223> Description of Artificial Sequence: Primer

<400> 22

gcaagcttaa tgtgacggtg aaataataac gg

32

<210> 23

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: HindIII Site

<400> 23

aagctt

6

<210> 24

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 24

taggtacctg gcaggtaaag actctgctc

29

<210> 25

<211> 6

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: KpnI Site

<400> 25

ggtacc

6

INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 00/00988

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/82 A01H5/10 A01H5/00 C12N15/29

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, EPO-Internal, BIOSIS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 18948 A (JAIN RAVINDER KUMAR ;THOMPSON ROBERTA GAIL (CA); CANADA NAT RES CO) 7 May 1998 (1998-05-07) cited in the application the whole document	1-3,6-9, 12,13
X	SAMMOUR REDA HELMY: "Proteins of linseed (Linum usitatissimum L.), extraction and characterization by electrophoresis." BOTANICAL BULLETIN OF ACADEMIA SINICA (TAIPEI), vol. 40, no. 2, April 1999 (1999-04), pages 121-126, XP000945031 ISSN: 0006-8063 the whole document	1-3,6-9, 12,13
	-/-	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

15 December 2000

Date of mailing of the international search report

05/01/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Bilang, J

INTERNATIONAL SEARCH REPORT

Inter national Application No

PCT/CA 00/00988

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>BORGMEYER J R ET AL: "ISOLATION AND CHARACTERIZATION OF A 25 KDA ANTIFUNGAL PROTEIN FROM FLAX SEEDS"</p> <p>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS,</p> <p>vol. 187, no. 1, 1992, pages 480-487,</p> <p>XP000941778</p> <p>ISSN: 0006-291X</p> <p>the whole document _____</p>	<p>1-3,6-9,</p> <p>12,13</p>

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/CA 00/00988

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9818948 A	07-05-1998	AU 4858897 A	22-05-1998
		GB 2332908 A	07-07-1999